University of Alberta

Accurate Estimation of Standard Errors in the Global Averaging of Surface Temperatures

by

Guilong Li

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

in

Applied Mathematics

Department of Mathematical Sciences

Edmonton, Alberta

Spring 2001

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-60454-3

Canada

University of Alberta

Library Release Form

Name of Author: Guilong Li

Title of Thesis: Accurate Estimation of Standard Errors in the Global Averaging of Surface Temperatures

Degree: Master of Science

Year This Degree Granted: 2001

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Guilong Li

Department of Mathematical Sciences
University of Alberta
Edmonton, Alberta
Canada. T6G 2G1

December 5, 2000

UNIVERSITY OF ALBERTA

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Accurate Estimation of Standard Errors in the Global Averaging of Surface Temperatures** submitted by **Mr. Guilong Li** in partial fulfillment of the requirements for the degree of **Master of Science** in Applied Mathematics.

Dr. Samuel Shen (Supervisor)

Dr. Peter Hooper (Chair)

Dr. Edward P. Lozowski
(Earth and Atmospheric Sciences)

Date: November 23, 2000

Abstract

This thesis uses an optimal statistical method for averaging global surface temperatures. The method minimizes the standard error budget resulting from statistical sampling errors due to station gaps and random data errors due to instrumental and human factors. Empirical orthogonal functions are used to represent the inhomogeneous covariance structure of the temperature field, and are computed from the following datasets: the Jones' land and United Kingdom Meteorological Office (UKMO)'s ocean 5°×5° data (1949-1998), the National Center for Environmental Prediction/National Center for Atmospheric Research Reanalysis data (1949-1998), the Climate Prediction Center's optimally interpolated sea surface temperature data (1982-1999), and the National Climatic Data Center's blended data from Global Historical Climatology Network and Special Sensor Microwave Imager (1992-1999). The Jones' box-data (1856-1998) over the land and UKMO's box-data (1856-1998) over the ocean are used as the observations, which are optimally averaged by our optimal averaging method. Using this method, one can generate not only the global average surface temperature of the monthly mean, annual mean, or decadal mean, but also an estimate of the error in the averaging process. The errors resulting from our method are only about 30% of those in the earlier assessments of the same quantities. The behavior of global averages is dominated by the sea surface temperature, whose spatial average has smaller standard errors compared with land data. Global warming trends during 1920-1944 and 1978-1998 are obvious. The standard error decreases from 0.065°C prior to 1900 to 0.03°C in recent time. These errors mainly result from random errors in the observations rather than sampling errors due to station gaps.

Acknowledgement

The results of this work were obtained during my Master's degree studies in the Department of Mathematical Sciences, University of Alberta. I would like to express deep gratitude to my supervisor Professor Samuel Shen, whose guidance and support were crucial for the successful completion of this project. I also want to thank Alan N. Basist, National Climatic Data Center, National Oceanic and Atmospheric Administration. We had many discussions on this topic and he gave me much inspiration on this project. Basist provided the weaver program which helped our data analysis.

Professor P.D. Jones of the University of East Anglia, the United Kingdom, kindly permitted me to use his land and ocean data and their error results. The author also wishes to thank the Hadley Center for Climate Prediction and Research, the United Kingdom. They provided the preprints of some of their recent work, which gave me a complete perspective on my own results.

The Department of Mathematical Sciences, University of Alberta, offered me the opportunity of a Teaching Assistantship that supported my study toward this degree. I also wish to thank the following: Darren Griffith, Hongtao Yang and Jeongyup Lee, and finally Marion Benedict who is always available whenever help is needed.

Contents

1 Introduction **1**

2 Data **8**

3 EOF computing **13**

3.1 The definition of continuous and discrete EOF 14

3.2 Computing EOF when $Y > J$ 16

3.3 Computing the spatial EOF from the temporal EOF when $Y < J$ . . . 17

4 The optimal averaging method and its MSE **20**

5 Optimal averaging procedure **25**

6 Optimal averaging results **29**

6.1 Eigenvalues and their validation 29

6.2 Eigenfunctions 31

6.3 Optimal averaging results 33

6.4 Error estimation 37

6.5 Weight distribution 38

7 Conclusion and Discussion **42**

7.1 Stability of EOF 42

7.2 Comparison between optimal averaging and spatial interpolation 45

7.3 Conclusions 45

List of Figures

Fig. 1 Comparison among monthly averaged temperature field for January 1983: (a) Jones data (b) Reanalysis data (c) Weaver data.

Fig. 2 Same to the Fig. 1 but comparison among (a) Jones data (b) Reanalysis data (c) smoothed weaver result.

Fig. 3 Number of data boxes in January from 1856 to 1998.

Fig. 4a The global anomaly temperature distribution in January 1900.

Fig. 4b The global anomaly temperature distribution in January 1983.

Fig. 5 The area-weight distribution for the Jones data in January 1983.

Fig. 6 The eigenvalues for each mode. The OI+GHCN blended data were used to do the EOF analysis.

Fig. 7a First 6 eigenfunctions in January for the OI+GHCN blended data.

Fig. 7b First 6 eigenfunctions in July for the OI+GHCN blended data.

Fig. 8a First 6 eigenfunctions in January for the weaver data.

Fig. 8b First 6 eigenfunctions in July for the weaver data.

Fig. 9 The OA result comparison between all modes and 12 modes, solid line for all modes and dash line for 12 modes. (a) Annual average (b) 10-year running average (c) Difference between all modes and 12 modes.

Fig. 10 Same as Fig. 9, but for a comparison between OA and Jones' area-weighted average (OI+GHCN blended data as EOF analysis for OA).

Fig. 11 Same as Fig. 10, but weaver data as EOF analysis for OA.

Fig. 12 Sampling and observational error using the OI+GHCN blended data for OA analysis.

Fig. 13 Same as Fig. 12 but using the weaver data for OA analysis.

Fig. 14 Sampling and observational error using Jones' area-weighted average.

Fig. 15a The weight distribution in January 1900 (OI+GHCN blended data for EOF analysis).

Fig. 15b The weight distribution in January 1983 (OI+GHCN blended data for EOF analysis).

Fig. 16a The weight distribution in January 1900 (Weaver data for EOF analysis).

Fig. 16b The weight distribution in January 1983 (Weaver data for EOF analysis).

Fig. 17a The weighted temperature in January 1900.

Fig. 17b The weighted temperature in January 1983.

Fig. 18 Optimal regional average of the anomaly temperature for 30-degree width bands.

Fig. 19a First six modes for 1949-1973 Reanalysis data EOF analysis.

Fig. 19b The 7-12th modes for 1949-1973 Reanalysis data EOF analysis.

Fig. 20a First six modes for 1974-1998 Reanalysis data EOF analysis.

Fig. 20b The 7-12th modes for 1974-1998 Reanalysis data EOF analysis.

Fig. 21a First six modes for 1949-1998 Reanalysis data EOF analysis.

Fig. 21b The 7-12th modes for 1949-1998 Reanalysis data EOF analysis.

Fig. 22 Comparison between first 25-year and second 25-year OA analysis. (a) Annual average (b) 10-year running average (c) Difference between the two optimal averagings

Fig. 23 Comparison between 50-year and 25-year OA analysis. (a) Annual average (b) 10-year running average (c) The difference between the optimal averaging (total period) and two half period OAs

List of acronyms

CPC - Climate Prediction Center

ECMWF - Medium-Range Weather Forecasts

ENSO - El Niño-Southern Oscillation

EOF - Empirical Orthogonal Functions

GCM - General Circulation Model

GFDL - Geophysical Fluid Dynamics Laboratory

GHCN - Global Historical Climate Network

IPCC - Intergovernmental Panel on Climate Change

MSE - Mean Square Error

NAO - North Atlantic Oscillation

OA - Optimal Averaging

OI - Optimal Interpolation

RMSD - root-mean-square difference

UKMO - United Kingdom Meteorological Office

SST - Sea Surface Temperature

Chapter 1

Introduction

Because of the overwhelming concern about climate change, an accurate estimation of global average surface temperature for past climate is of great importance. According to IPCC 1995 report, the observed decadal temperature changes are about 0.3-0.6°C since the mid-19th century, indicating that care must be taken to reduce the uncertainty associated with this estimation. The uncertainty includes mainly the random observational data errors and sampling errors of the temperature anomaly. Without systematic estimation, the combination of the two errors was assessed to be around 0.15°C, which is about 25-50% of the warming amplitude since the mid-19th century (IPCC, 1995).

Up to now, there is less than 300 years of instrumental data. The globally well-covered data have only about 150 years history. Climate from before the thermometer era must be deduced from paleoclimatic proxy records. These include tree-rings, pollen series, faunal and floral abundances in deep sea cores, isotope analysis from coral and ice cores, and diaries and other documentary evidence. Cook et al. (1994) used two alternative spatial regression methods to reconstruct climate from tree-rings: orthog-

onal spatial regression and canonical regression. The two methods have a common foundation in least-squares theory. The results showed that there was little difference between orthogonal spatial regression and canonical regression with regard to reconstructing climate from tree-rings. Each method was equally good (or deficient) in its ability to reconstruct temperature using different tree-ring data bases. This strong similarity of results must be due to the fact that each method is based on least-squares theory. The failure of each method to reconstruct some regions of each climatic grid shows that these sophisticated statistical procedures cannot extract a climatic signal from tree-rings when it is very weak or effectively non-existent in the data.

Because the instrumental data distribution cannot cover every point in the world, instrumental data has a problem of sampling gaps. Whether can we use less data in the early period to get an accurate global average temperature? Many methods have been developed on studying the incomplete spatial sampling.

Oort (1978) used the output from a Geophysical Fluid Dynamics Laboratory (GFDL) global general circulation model (GCM) to test the adequacy of an 855-station rawinsonde network for global circulation studies. He compared globally averaged temperature based on numerical output from the model at its full spatial resolution with similar temperature values determined only at grid points near actual rawinsonde locations. He pointed out that the errors were mainly coming from the spatial gaps. Trenberth and Olson (1991) used nine years of global European Centre for Medium-Range Weather Forecasts (ECMWF) analyses to test the representativeness of a 63-station rawinsonde network. Like the Oort study, they determined differences between globally averaged anomalies sampled at all available grid points of the ECMWF analyses and those sampled at only the 63 grid points nearest to the stations. They found the root-mean-square difference (RMSD) averaged over all months and seasons to be

0.18°C and 0.15°C, respectively.

Jones et al. (1986a,b) used the "frozen grid" approach. Their calculation of the hemispheric average temperature for all years used only those grid points that operate 80% of the time during a particular decade, to estimate the effects of incomplete spatial sampling. They compared annual and hemispherically averaged temperature anomalies, determined using either all of the observed grid points available at a given time or frozen grids, that is, grids made up of only the observed grid points available during an earlier period. For the frozen grids available after the mid-19th century, the differences were typically a few tenths of a degree centigrade for both hemispheres. Hansen and Lebedeff (1987) made a comparison somewhat similar to that of Jones et al. (1986a,b) to test the effects of poor spatial sampling, but they used output from a 100-year GCM run, rather than observed data. In their comparison, it is important that the models exhibit spatial variations that are similar to those of the atmosphere. Hansen et al. (1987) showed that available comparisons indicate that the temperature variability on large spatial scales in their model is comparable in magnitude to the variability in the real world. They compared globally averaged temperature anomalies based on the "true" model values with those determined from only those grid points that were near available stations during each decade. For annual averages, they found an RMSD of 0.07°C for station locations available in the 1880s and 0.02°C for those available in the 1960s.

Madden et al. (1993) used model data to study the effect of imperfect spatial and temporal sampling on estimates of the global mean temperature. Two methods used in his study were the simple statistical model and the Monte Carlo method. The simple model method shows that when a single adjustable parameter is suitably chosen, a simple statistical model based on the theory for a randomly sampled finite population

can give values reasonably close to the empirical estimates. The RMSD between perfectly and imperfectly spatially sampled temperatures varied from 0.224°C before the turn of the century to 0.045°C after 1950. The Monte Carlo method is based on the sampling of synthesized maps that have the same spatial spectrum and are reasonably close to the empirical ones. He also pointed out that if there are reasonable estimates of the spatial variance of the number of independent estimates in the global distribution of actual surface temperature data, or, alternatively, their spatial spectrum, one might be able to make first-order estimates of the spatial sampling errors. The only thing one needs is the temporal variance of the temperature to make best- or worst-case scenario estimates of the error due to imperfect temporal sampling.

In order to show how many stations would be required to reasonably estimate the global average temperature, Jones (1994) analyzed the station monthly mean air temperature series, with over 1000 more stations than previously used and employing a $5° \times 5°$ box resolution with 1961-1990 climatology. Even if he had put 1088 additional stations, the hemispheric-mean temperature series would have achieved similar results with fewer stations. This result confirms the robustness of hemispheric-mean temperature with fewer stations. Jones et al. (1997) also made a systematic study of the standard errors and found that the annual errors on an interannual timescale decreases from 0.091°C prior to 1900 to 0.059°C since 1951. Thus, for the standard error calculation, increasing the density of stations over land areas and the number of measurements per month for ocean squares will reduce the error. The substantial uncertainty in the assessment undermines the significance claim of global warming. It gives a high noise level in the climate change signal and hence affects the detectability of global warming.

Several methods are introduced in this work to reduce standard error estimation

due to imperfect sampling and instrumental errors. The question is which method is the best for estimating global temperature variation. The optimal averaging (OA) method used here belongs to the category of objective analysis first invented in the former Soviet Union (Kagan, 1979 (English edition 1997), Vinnikov, 1990). The method uses a covariance model to achieve one spatial average of a field. If the spatial covariance structure of the temperature is known, the OA should be a closer estimate to the true mean than a simple area-weighted average. Additionally, it provides an intrinsic estimate of the error due to imperfect spatial sampling. In earlier times, covariance models were typically assumed to be simple and homogeneous. Weber and Madden (1995) used the correlation decay-length to do the OA to determine the global average temperature. The method considered the spatial inhomogeneity of the temperature field. In their study, they divided the globe into six latitudinal bands of 30° width. The covariance structure based on the correlation model for each band was calculated. Their results showed that the OA procedures provide a better approximation to the true mean than a simple area-weighted average.

However, this method did not consider the systematic global structure, such as El Niño-Southern Oscillation (ENSO), North Atlantic Oscillation (NAO) etc. The atmospheric circulation is the main control behind regional changes in wind, temperature, precipitation, soil moisture and other climatic variables. Variations in many of these variables are strongly related through large-scale features of the atmospheric circulation, as well as through interactions involving the land and ocean surfaces. Two well-known examples of such large-scale features are the ENSO and the NAO, both of which are closely related to climatic fluctuations in many regions. Evidence of associated changes or variations in the atmospheric circulation may enhance confidence in the reality of observed changes and variations in the climate variables in these areas.

Shen et al. (1994) developed an OA method and applied it to inhomogeneous fields using empirical orthogonal functions (EOF) for global OA. The spatial inhomogeneity of the covariance function is resolved by the superposition of the EOF model. Their OA results showed that the OA has the smallest sampling error compared to the other methods. But random error was not considered until a later study (Shen et al., 1998). Kim et al. (1996) used this idea to estimate the spherical harmonic coefficients on a sparse network of observational data. The global average temperature and sampling error were calculated. The OA performed best compared to the uniform averaging and Gaussian averaging of the data. Kaplan et al. (1997) developed a method to reconstruct the Atlantic sea surface temperature. The method filled spatial gaps in the data and minimized sampling error. Similar to Shen and Kim's result, the advantage of Kaplan's method, inherited from the classical least-squares approach, is its ability to provide error estimates for analyzed data. Further research on optimal regional averaging methods was undertaken by Shen et al. (1998). This method was tested on tropical Pacific SST data and random observational error was considered. Their idea was that an inhomogeneous covariance function could be represented by its eigenvalues (i.e. variance) and eigenfunctions (i.e. EOF).

In our present work, we attempt to bring the previous researches together to study the OA of global temperature. The purpose of this thesis is to show that the OA method can be used in such a way that the uncertainty in assessing global warming can be reduced significantly. This work follows the theory by Shen et al. (1998) and optimally averages the global anomaly temperature. The main idea is to minimize the mean square error (MSE) and to obtain the optimal weights for each data box. Not only the sampling error but also random error for global average temperature will be considered. Then with these optimal weights, the OA temperature can be easily

calculated.

The thesis structure is as follows. Section 2 will explain the data used in the study. Sections 3, 4 and 5 will explain the OA theory. In Section 6, numerical results are described and standard errors are assessed. Section 7 summarizes the main conclusions and gives some discussion.

Chapter 2

Data

Several datasets are used in this study, divided into two groups. The first group contains the recent data in all 5° × 5° boxes used to extract the statistical EOF and variance, which will be used by optimal averaging. The second group includes the historical data only in some 5° × 5° boxes, used to calculate the global average temperature. In order to compare the effect of different EOF on OA, two datasets are used in group one. One comes from Climate Prediction Center's (CPC) Optimal Interpolation (OI) data and Global Historical Climate Network (GHCN) blended data. Another spans from the 1949 to 1998 Jones land data and the United Kingdom Meteorological Office (UKMO) SST data. In group two are the Jones land data and UKMO ocean data from 1856 to 1998. The 1961-1990 climatology is computed separately for each dataset and is used to compute the climate anomalies. The climatology is 1961-1990 averages rather than the 1951-1980 period, because the new reference period contains more SST data and more data from currently operating land stations. Jones' (1997) interannual error is introduced as the observational error. This error is used in calculating optimal weights and estimating the total error. Each dataset consists of the

monthly anomaly temperature on a 5° × 5° longitude-latitude grid box. If there were data in each box, the total number of boxes would be 72 × 36. The center of first box is at 177.5°W,87.5°S.

The OI SST analysis was produced weekly on a one-degree latitude-longitude grid. The analysis used in situ and satellite SST plus sea-ice cover. Before the analysis was computed, the satellite data were adjusted for biases using the method of Reynolds (1988) and Reynolds and Marsico (1993) in order to improve the large scale accuracy. A description of the OI analysis can be found in Reynolds and Smith (1994), and examples of the effect of recent corrections can be found in Reynolds (1993). Because the OA method requires 5° × 5° monthly data, one can adjust the 1° × 1° monthly OI data to 5° ×5° monthly OI data by simply average the grid data in the 5° × 5° box. As OI data cover only ocean areas (from January 1982 to June 1999), a combination of the OI data with the GHCN blended data on land (from January 1992 to June 1999) was used. Both anomalies were with respect to 1961-1990 climatology. The blended data anomalies were obtained using the climatologies of Reanalysis in the periods of 1992-1998 and 1961-1990. The offset between the two base periods was removed from the GHCN blended data anomalies, whose base period was 1992-1998. For the OI data, the boxes over land are empty. For the GHCN blended data, the boxes over ocean and snow-covered areas are likewise empty. These empty boxes are then filled in by the Reanalysis data. The variances of the Reanalysis data are smaller than the true variances in the tropical areas and larger than the true ones in the high-latitude areas. Such properties limit the utility of the Reanalysis data in computing the EOF.

The 1856-1998 dataset is a combination of the land temperature anomalies described by Jones (1994) and the UKMO SST data. The Jones land data are based upon station observations, simply taken as an average over each 5°×5° box. On the

other hand, the UKMO ocean data are based on ship observations. Various corrections were made before the two datasets were merged. For example, a bias correction was applied to the SST anomalies before 1942 to account for changes from bucket-base to intake-base temperature measurements (Folland and Parker 1995). In the remainder of this paper, these merged data are simply referred to as "Jones data". Now the Jones data has missing data in $5°\times5°$ boxes. Namely, there are "holes" in the dataset. These "holes" will be filled by the Reanalysis data. The data from the last 50 years, 1949-1998, are taken as a second dataset in Group 1 for EOF analysis. There are two reasons to select the period from 1949 to 1998 for EOF analysis. First, since the Reanalysis data spans the period 1949 to 1998, it can be used to fill in the "holes" in the Jones data. Another reason is that the sample density of Jones data becomes large after the Second World War. But there is a further problem of whether we can simply merge two datasets, because they are two different datasets and have different strengths and weakness. The Reanalysis data reflects large scale patterns well, but are too smooth at low latitude and have too large variance at high latitude. Jones dataset reflect a more realistic field but have missing data. Is it possible to combine the advantages of the two datasets? One method, called weaver, is applied. This method can be explained most clearly by using figures. Fig. 1(a) and (b) are the Jones data and the Reanalysis data in January 1983, respectively. One may observe that the large scale patterns of the Reanalysis data and the Jones data are similar, but that there are differences in details between the two datasets. For example, in Fig. 1, there is a strong ENSO anomaly in the Eastern Pacific Ocean in the Jones data, but the corresponding anomaly is weaker and more smooth in the Reanalysis data. One can also notice the shapes of anomalies in Russia are the same, but the strength of the anomaly is different. The weaver method uses the Reanalysis data

pattern and the Jones data strength to form a more realistic dataset. First, it takes the difference of the two datasets, and then disperses the difference value in each box to the neighboring points to get a new difference dataset. Then the new difference dataset is added to the Reanalysis data to form the weaver data, as shown in Fig. 1(c). Thus in Fig. 1(c), it can be seen that the weaver data keeps the information of the Jones data and the pattern of the Reanalysis data. The Jones data bring some small scales into the Reanalysis data. These kinds of small structures will influence the EOF analysis. One may use a spatial five-point average to remove the small scales. Fig. 2(c) shows the averaged result. Compared to Fig. 1(c), it is clear that the small scales have been removed. Thus, a new dataset whose structure is taken from the Reanalysis and whose amplitude is taken from the Jones data has been developed.

Jones et al. (1997) developed a method to calculate the standard errors of 5°×5° boxes for any regional/hemispheric/global time series of a climate variable. The main idea is that the standard error depends on station numbers (n) and average the intersite correlation ($\bar{r}$) between the n stations in the box. The formula is written as $SE^2 = \bar{s}_i^2 \bar{r}(1-\bar{r})/[1+(n-1)\bar{r}]$, where $\bar{s}_i$ is the characteristic variance of a station time series within the grid box. Using this method, they obtained typical standard errors, estimated for annual data as 0.059°C since 1951. Prior to 1900, the standard error was 0.091°C.

A time series of the number of 5°×5° data boxes sampled in January is shown in Fig. 3. The sparse samples occurred either in the first few decades from 1856 or during the First and Second World Wars. The sampling distribution in January 1900 (Fig. 4a) shows that there were almost no data poleward of 70° latitude, and there were often large numbers of "holes" over non-industrialized land regions at the beginning of the twentieth century and earlier. Over the oceans, the Atlantic tends

to be the best sampled and the Pacific the worst sampled. The southern ocean is also poorly sampled. In the 1980s', the sampling distribution was more dense than at the beginning of the twentieth century. Fig. 4b shows the sampling distribution in January 1983. Most of the area except the poles were well sampled.

The data used in this work can be summarized in table 1.

Data comparison

Name	Temporal resolution	Spatial resolution	Period	Property
Reanalysis	1 month	5° × 5°	January 1949 - December 1998	Globe
Optimal Interpolation	1 month	1° × 1°	January 1982 - June 1999	Ocean
GHCN blended	1 month	5° × 5°	January 1992 - June 1999	Land
Jones	1 month	5° × 5°	January 1856 - December 1998	Globe with "holes"

Chapter 3

EOF computing

EOF is a popular method for analyzing the inhomogeneous covariance structure of a field. It is very useful in data compression and physical pattern analysis. In meteorology, one can use a few modes to reconstruct a meteorological filed, such as temperature or precipitation. As instrumental records extended back only to the mid-19th century and the data were sparse in the early period, as shown in Fig. 4a, it is necessary to use various proxy approaches to derive the history of temperature. In addition to historical records, proxy measurements such as corals, tree-rings, ice cores, and varve sediments have also been employed to derive the temporal history of temperature variability.

Several methods can be used in the calculation of global average temperature. The simplest method, whose accuracy is in question, is to average all the boxes which contain data. One problem with this method is that the area of a $5°\times5°$ box varies according to latitude. A larger area should make a larger contribution to the global temperature average compared to a smaller one. For example, when selecting one box in the tropics and one at mid-high latitude (60°N), the areas for these two boxes are 309,778 km^2 and 154,889 km^2, respectively. This means that at 60° south and

north latitudes, the area in one box is half the area of a box in the tropics. This is shown clearly in Fig. 5. So the area influence must be considered. This leads us to the area-weighted average method which considers area influence (Jones et al., 1999). But this causes another problem. Because the anomaly temperature is inhomogeneous over the globe, the area-weighted average does not consider the anomaly temperature's inhomogeniety, since in this method the same latitude has the same area-weight in the area-weighted average. The temperature anomaly is warmer than average during *El Niño* and cooler than average during *La Niña* in the Eastern tropical Pacific (Fig. 4b). Inhomogenieties also exist in the mid-latitude regions, where different patterns occur along the same latitude band. The homogeneous area-weighted average has certainly neglected the inhomogeneous physical properties. In order to reflect the inhomogeniety, the EOF method was introduced. Adapted to the data's inhomogeniety and sparse distribution, this optimal method to the global average temperature was developed by Shen et al. (1994). Subsequently, random observational errors were included and regional temperature averaging was studied (Shen et al. 1998).

3.1 The definition of continuous and discrete EOF

The covariance function is defined by

$$\rho(\hat{\mathbf{r}}, \hat{\mathbf{r}}') = \langle T(\hat{\mathbf{r}}, t) T(\hat{\mathbf{r}}', t) \rangle. \tag{3.1}$$

The continuous EOF $\psi_n(\hat{\mathbf{r}})$ are the eigenfunctions of the covariance function defined by

$$\int_\Omega \rho(\hat{\mathbf{r}}, \hat{\mathbf{r}}') \psi_n(\hat{\mathbf{r}}')\, d\Omega' = \lambda_n \psi_n(\hat{\mathbf{r}}), \quad n = 1, 2, 3, \cdots, \tag{3.2}$$

where Ω is the domain, λ_n is the eigenvalue which is equal to the variance of the temperature's projection on the n-th EOF:

$$\lambda_n = \langle \left(\int_\Omega T(\hat{\mathbf{r}}, t)\psi_n(\hat{\mathbf{r}})\, d\Omega \right)^2 \rangle. \tag{3.3}$$

The EOF have orthogonality and normalization properties

$$\int_\Omega \psi_m(\hat{\mathbf{r}})\psi_n(\hat{\mathbf{r}}) d\Omega = \delta_{mn}, \tag{3.4}$$

which is equal to unity if $m = n$, and zero otherwise.

The region Ω is divided into J grid boxes. The discretization of the above integral on these J boxes is

$$\sum_{j=1}^{J} \rho(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j)\psi_n(\hat{\mathbf{r}}_j)A_j = \hat{\lambda}_n\psi_n(\hat{\mathbf{r}}_i), \quad i = 1, 2, \cdots, J, \tag{3.5}$$

where A_j is the area of the grid box j, and $\hat{\lambda}_n$ is the approximation of λ_n due to discretization. In the case of $5^\circ \times 5^\circ$ grid boxes, the box area is

$$A_j = R^2 \times \left(\frac{5}{180}\pi\right) \times \left(\frac{5}{180}\pi\right) \cos\phi_j, \tag{3.6}$$

where ϕ_j is the latitude of the center of the box j, and R is the radius of Earth, approximately 6,376 km.

To make the equation (3.5) symmetric, it can be rewritten as

$$\sum_{j=1}^{J} \left(\sqrt{A_i}\rho(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j)\sqrt{A_j}\right) \left(\psi_n(\hat{\mathbf{r}}_j)\sqrt{A_j}\right) = \hat{\lambda}_n \left(\psi_n(\hat{\mathbf{r}}_i)\sqrt{A_i}\right), \quad i = 1, 2, \cdots, J, \tag{3.7}$$

The new discrete eigenvalue problem is then

$$\sum_{j=1}^{J} \hat{\rho}(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j)\hat{v}_j^{(n)} = \hat{\lambda}_k^{(n)}\hat{v}_i^{(n)}, \tag{3.8}$$

where

$$\hat{\rho}(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j) = \sqrt{A_i}\ \rho(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j)\sqrt{A_j} \tag{3.9}$$

is the area-modified covariance matrix, and

$$\hat{v}_j^{(n)} = \psi_n(\hat{\mathbf{r}}_j)\sqrt{A_j} \tag{3.10}$$

is the area-modified eigenfunction, which is defined as the discrete EOF. They satisfy the normalization condition

$$\sum_{j=1}^{J} \left(\hat{v}_j^{(n)}\right)^2 = 1. \tag{3.11}$$

Therefore, based upon equation (3.8), one can obtain the eigenvalues λ_n and the EOF

$$\psi_n(\hat{\mathbf{r}}_j) = \frac{\hat{v}_j^{(n)}}{\sqrt{A_j}}. \tag{3.12}$$

The length of observations is denoted by Y. Depending on the values of Y and J, the calculations of the eigenvalues and eigenfunctions follow different algorithms since a covariance matrix computed from insufficient observations is not of full rank. The next two sub-sections describe these algorithms.

3.2 Computing EOF when $Y > J$

To compute the EOF for each month, two anomaly datasets are used: OI+GHCN blended+Reanalysis data and weaver data (1949-1998).

These data are put into one matrix of order $J \times Y$, which is denoted by $\mathbf{T}$:

$$\mathbf{T} = [T(\hat{\mathbf{r}}_j, t)], \quad j = 1, 2, 3, \cdots, J, \quad \text{and} \quad t = 1, 2, 3, \cdots, Y, \tag{3.13}$$

where J is the total number of boxes in the domain Ω, and Y is the total number of years of the data. For example, for the January EOF, the combination of the OI (1982-1999) and GHCN blended (1992-1999) gives a total of 26 years of data, and hence $Y = 26$.

The spatial covariance matrix $[\rho(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j)]$ is

$$\mathbf{TT}'/Y, \tag{3.14}$$

which is a $J \times J$ order matrix, where $'$ indicates the transpose of the matrix. The maximum number of independent column vectors in this matrix is Y. When the data stream is sufficiently long such that $Y > J$, the spatial covariance matrix is most likely full rank, and one can compute the discrete EOF directly from (3.7). The discretized eigenvalue problem (3.7) from the integral eigenvalue problem requires the inclusion of an area-factor matrix

$$\sqrt{\mathbf{A}} = [\sqrt{A_i}\delta_{ij}], \tag{3.15}$$

where $\sqrt{A_i}$ is the area of the grid box i, and the diagonal matrix is of order $J \times J$. The discrete spatial eigenvalue problem with the inclusion of the area factor is then

$$[\hat{\rho}]\hat{v} = \hat{\lambda}\hat{v}, \tag{3.16}$$

where the spatial covariance matrix with the area factor is

$$[\hat{\rho}] = \frac{1}{Y}(\sqrt{\mathbf{A}}\mathbf{T})(\sqrt{\mathbf{A}}\mathbf{T})'. \tag{3.17}$$

3.3 Computing the spatial EOF from the temporal EOF when $Y < J$

Often, the dataset used for computing the covariance matrix has not been sufficiently long, meaning that the spatial covariance matrix is not of full rank. In such a case, the spatial EOF can be computed indirectly from the temporal EOF.

Because each month has a different physical pattern, the EOF has to be computed according to month. For example, if one works on the globe with $5° \times 5°$ boxes, then

$J = 36 \times 72 = 2592$. For January, the OI data (January 1982 to June 1999) and GHCN blended data (January 1992 to June 1999) can provide only 26 years of data, and the index t in the above matrix runs from 1 to 18 for the OI data, and from 19 to 26 for the GHCN blended data, so $Y = 26$. Hence, $Y < J$, the spatial covariance matrix is not of full rank, and its determinant vanishes. However, one can exchange space and time by transposing the data matrix first:

$$\mathbf{T}'\mathbf{T}/J. \tag{3.18}$$

This is the temporal covariance matrix, whose order is $Y \times Y$.

The temporal eigenvalue problem with the inclusion of the area factor is

$$[\hat{\rho}_t]\mathbf{u} = \tilde{\lambda}\mathbf{u}, \tag{3.19}$$

where the temporal covariance matrix with the area factor is

$$[\hat{\rho}_t] = \frac{1}{J}(\sqrt{\mathbf{A}}\mathbf{T})'(\sqrt{\mathbf{A}}\mathbf{T}), \tag{3.20}$$

and the eigenfunction $\mathbf{u}$ is normalized:

$$\sum_{k=1}^{Y} u_k^2 = 1. \tag{3.21}$$

Multiplying the temporal eigenvalue problem by the matrix $(J/Y)(\sqrt{\mathbf{A}}\mathbf{T})$ and comparing the resulting equation with the spatial eigenvalue problem (3.16) yields the following relationship:

$$\hat{\lambda} = \frac{J}{Y}\tilde{\lambda} \tag{3.22}$$

$$\hat{v} = \frac{1}{\sqrt{\tilde{\lambda}}}\sqrt{\mathbf{A}}\mathbf{T}\mathbf{u}. \tag{3.23}$$

This vector $\hat{v}$ is a unit vector because

$$\hat{v}'\hat{v} = \frac{1}{\sqrt{\tilde{\lambda}}}\mathbf{u}'(\sqrt{\mathbf{A}}\mathbf{T})'\frac{1}{\sqrt{\tilde{\lambda}}}\sqrt{\mathbf{A}}\mathbf{T}\mathbf{u} = \mathbf{u}'\mathbf{u} = 1.$$

Finally, the EOF and the corresponding eigenvalues are

$$\psi_n(\hat{\mathbf{r}}_j) = \frac{\hat{v}_j^{(n)}}{\sqrt{A_j}}, \quad j = 1, 2, \cdots, J, \quad \text{and} \quad n = 1, 2, \cdots, \tag{3.24}$$

$$\lambda_n = \hat{\lambda}_n. \tag{3.25}$$

There are two benefits derived through space and time exchange. First, the problem of determinant vanishing does not exist. Second, because the dimension of $Y \times Y$ is less than that of $J \times J$, the EOF analysis will require less calculation through this transposing.

Chapter 4

The optimal averaging method and its MSE

The average of the field T over a region Ω is

$$\bar{T} = \frac{1}{A}\int_{\Omega} T(\hat{\mathbf{r}})\, d\Omega \tag{4.1}$$

where A is the area of the region Ω. Our objective is to use the sampling data $\tilde{T}(\hat{\mathbf{r}}_i)$ to estimate this quantity with maximum accuracy. The method employed here is referred to as optimal averaging (OA) and has been discussed by Kagan (1979), Vinnikov et al. (1990), Smith et al. (1994), Kim et al. (1996), Kaplan et al. (1997), and Shen et al. (1994 and 1998). Here the version of OA is similar to that of Shen et al. (1998) in which EOF defined on the entire globe were computed for 26 (24) years from the OI+GHCN blended data and from 50 years of weaver data (1949-1998). These two datasets are used in the series expansion. The method of Shen et al. (1998) is extended in this study from a regional average to a global average. In this work, however, unlike the spherical harmonics method in Kim et al. (1996), an area factor will be taken into account when computing eigenvalues and EOF. Namely, the EOF are computed

from area-weighted OI data for the period of January 1982 to June 1999 and GHCN blended data from January 1992 to June 1999 and for the weaver of the Jones and the Reanalysis data. Because Jones (1999) calculated an area-weighted average for each month, the OA of each month will be calculated for comparison. The annual average will be obtained through averaging the results of 12 months. The effect of random observational errors is included in the total error.

The linear estimator of the average, denoted by $\hat{\bar{T}}$, is

$$\hat{\bar{T}} = \sum_{j \in \mathcal{N}} w_j \tilde{T}_j, \tag{4.2}$$

where $\mathcal{N}$ denotes the observational network on which the gaped data are distributed and the weights w_j satisfy a normalization condition:

$$\sum_{j \in \mathcal{N}} w_j = 1. \tag{4.3}$$

This condition is needed because our data contain trends. So to guarantee that

$$\langle \hat{\bar{T}} \rangle = \sum_{j \in \mathcal{N}} w_j \langle \tilde{T}_j \rangle$$

when $\langle \tilde{T}_j \rangle = \langle \hat{\bar{T}} \rangle$, it is necessary to have the normalization condition (4.3) as discussed by Kagan (1979).

The sampling error is measured by the mean square error (MSE)

$$\epsilon^2 = \langle (\bar{\bar{T}} - \hat{\bar{T}})^2 \rangle, \tag{4.4}$$

and the covariance function is defined by

$$\rho(\hat{\mathbf{r}}, \hat{\mathbf{r}}') = \langle T(\hat{\mathbf{r}}) T(\hat{\mathbf{r}}') \rangle. \tag{4.5}$$

The following notations are adopted:

$$\rho_{ij} = \rho(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j) = \langle T_i T_j \rangle \tag{4.6}$$

and

$$\bar{\rho}_j = \frac{1}{A}\int_\Omega \langle T(\hat{\mathbf{r}})T_j\rangle d\Omega\,. \tag{4.7}$$

In terms of data $\tilde{T}_i$ and observation errors $\langle E_i^2\rangle$, the covariance matrix ρ_{ij} can be written as

$$\begin{aligned}\rho_{ij} &= \langle T_iT_j\rangle = \langle(\tilde{T}_i - E_i)(\tilde{T}_j - E_j)\rangle \\ &= \langle\tilde{T}_i\tilde{T}_j\rangle + \langle E_iE_j\rangle \\ &= \langle\tilde{T}_i\tilde{T}_j\rangle + \langle E_i^2\rangle\delta_{ij},\end{aligned} \tag{4.8}$$

where δ_{ij} is the Kronecker delta which is equal to unity when $i = j$ and zero otherwise.

The MSE can be written as

$$\begin{aligned}\epsilon^2 &= \langle(\bar{T} - \hat{\bar{T}})^2\rangle = \langle\bar{T}\bar{T}'\rangle - 2\langle\bar{T}\hat{\bar{T}}\rangle + \langle\hat{\bar{T}}\hat{\bar{T}}'\rangle \\ &= \frac{1}{A^2}\int_\Omega d\Omega\int_\Omega d\Omega'\langle T(\hat{r})T(\hat{r}')\rangle - \frac{2}{A}\langle\int_\Omega T(\hat{r})d\Omega\sum w_jT_j\rangle + \langle\sum_{i\in\mathcal{N}} w_i\tilde{T}_i\sum_{j\in\mathcal{N}} w_j\tilde{T}_j\rangle \\ &= \frac{1}{A^2}\int_\Omega d\Omega\int_\Omega d\Omega'\rho(\hat{\mathbf{r}},\hat{\mathbf{r}}') - 2\sum_{j\in\mathcal{N}} w_j\bar{\rho}_j + \sum_{i,j\in\mathcal{N}} w_iw_j\,\langle\tilde{T}_i\tilde{T}_j\rangle \\ &= \frac{1}{A^2}\int_\Omega d\Omega\int_\Omega d\Omega'\rho(\hat{\mathbf{r}},\hat{\mathbf{r}}') - 2\sum_{j\in\mathcal{N}} w_j\bar{\rho}_j + \sum_{i,j\in\mathcal{N}}{}' w_iw_j(\langle\bar{T}\bar{T}'\rangle + \langle E_i^2\rangle\delta_{ij}) \\ &= \frac{1}{A^2}\int_\Omega d\Omega\int_\Omega d\Omega'\rho(\hat{\mathbf{r}},\hat{\mathbf{r}}') - 2\sum_{j\in\mathcal{N}} w_j\bar{\rho}_j + \sum_{i,j\in\mathcal{N}} w_iw_j\,\rho_{ij} + \sum_{i\in\mathcal{N}} w_i^2\langle E_i^2\rangle.\end{aligned} \tag{4.9}$$

To minimize the MSE, a Lagrange function is constructed

$$\mathcal{L}[w_1,\cdots,w_N] = \epsilon^2(w_1,\cdots,w_N) + 2\Lambda\left(\sum_{j\in\mathcal{N}} w_j - 1\right) \tag{4.10}$$

where Λ is the Lagrange multiplier and N is the number of stations in the network $\mathcal{N}$. The partial derivatives

$$\frac{\partial \mathcal{L}}{\partial w_j} = 0, \qquad j = 1, 2, 3, \cdots, N,$$

and

$$\frac{\partial \mathcal{L}}{\partial \Lambda} = 0$$

lead to

$$\sum_{j \in \mathcal{N}} \rho_{ij} w_j + \langle E_i^2 \rangle w_i + \Lambda = \bar{\rho}_i, \quad i = 1, 2, \cdots, N; \tag{4.11}$$

$$\sum_{j \in \mathcal{N}} w_j = 1. \tag{4.12}$$

The solution of the above set of equations yields the optimal weights $w_1, w_2, \cdots, w_N$ for computing the OA by (4.2).

The covariance matrix $[\rho_{ij}]$ can be approximated by

$$\rho_{ij} = \frac{1}{M_\gamma} \sum_{\gamma=1}^{M_\gamma} \tilde{T}_i(\gamma) \tilde{T}_j(\gamma) + \langle \overline{E_i^2} \rangle \delta_{ij}. \tag{4.13}$$

In this expression it is assumed that the time series $T(\hat{\mathbf{r}}_i, t)$ satisfies an ergodic process (the ensemble average $\langle T(\hat{\mathbf{r}}_i, t) T(\hat{\mathbf{r}}_j, t) \rangle$ is equal to the temporal average), which is approximated by the above summation with respect to the time variable γ. M_γ is the maximal length of the data streams to be processed. It should be pointed out that $\tilde{T}_j(\gamma)$ may be serially correlated, but due to the short length of the data streams of the recent accurate observations, it is still the best approximation to estimate the covariance matrix $[\rho_{ij}]$ by (4.13) rather than throwing out some data so that the remaining data are serially independent of one another. From formula (4.13) one can see that the rank of the computed covariance matrix $[\rho_{ij}]$ is M_γ since M_γ is usually much less than the total number of grid boxes. Thus the covariance matrix $[\rho_{ij}]$ is often not a full rank matrix as discussed in the last section of Chapter 2.

To solve eqs. (4.11)- (4.12), one needs to find $\bar{\rho}_i$, the average of the covariance function around the station $\hat{\mathbf{r}}_i$. Hence, the original problem of averaging T is partially converted into an averaging problem of the covariance function. This conversion is important since it provides us a new way of evaluating sampling errors due to the fact that the $\bar{\rho}_i$ can be computed from the averages of the EOF. They can also be estimated from climate models, such as GCMs (Zwiers and Shen 1997) or even energy balance models (Kim and North 1993). The algorithms for computing EOF from the area-weighted data and for computing $\bar{\rho}_i$ are described in the next section.

Chapter 5

Optimal averaging procedure

Our method of computing $\bar{\rho}_i$ is illustrated using the monthly SST OI data from January 1982 to June 1999 and GHCN blended data from January 1992 to June 1999 and also the 50-year weaver data. Here these data are regarded as the "truth". The OI+GHCN blended anomalies are gridded on a $5^\circ \times 5^\circ$ grid, denoted by $\mathcal{OI}$. Let M_{OI} denote the length of the data stream (equal to 26 years for the first half year and 24 for the second half year of OI+GHCN blended data and 50 years of weaver data). Then eq. (4.13), with exclusion of the random error $\langle E_i^2 \rangle$, is used to compute the covariance matrix $[\rho_{ij}^{OI}]$ from the OI data. The exact eigenvalue problem is

$$\int_\Omega \rho(\hat{\mathbf{r}}, \hat{\mathbf{r}}')\psi_k(\hat{\mathbf{r}}')\, d\Omega' = \lambda_k \psi_k(\hat{\mathbf{r}}). \tag{5.1}$$

Here $\psi_k(\hat{\mathbf{r}})$ is the kth EOF (or mode) and λ_k is the variance (eigenvalue) of $T(\hat{\mathbf{r}})$ on the kth mode ($k = 1, 2, \cdots$). The approximate eigenvalues of the above continuum eigen problem can be estimated by a discretization procedure given by

$$\sum_{j \in \mathcal{OI}} \hat{\rho}_{ij} \hat{v}_j^{(k)} = \hat{\lambda}_k \hat{v}_i^{(k)}, \tag{5.2}$$

where

$$[\hat{\rho}_{ij}] = \left[\sqrt{A_i}\ \rho_{ij}^{OI} \sqrt{A_j}\ \right] \tag{5.3}$$

is the modified covariance matrix,

$$\hat{v}_j^{(k)} = \psi_k(\hat{\mathbf{r}}_j)\sqrt{A_j}, \tag{5.4}$$

are the modified eigenfunctions satisfying the normalization condition

$$\sum_{j \in \mathcal{OI}} \left(\hat{v}_j^{(k)}\right)^2 = 1, \tag{5.5}$$

and A_j is the area associated with the station $\hat{\mathbf{r}}_j$. For uniform latitude-longitude grid boxes, one has

$$A_j = \cos\phi_j \Delta\theta \Delta\phi \tag{5.6}$$

where ϕ_j is the latitude of $\hat{\mathbf{r}}_j$ and the $\Delta\theta$ and $\Delta\phi$ are the zonal and meridional box dimensions respectively, which are measured in radians. The linear spatial unit (i.e., the length unit) is in the radius of the Earth: $R = 6376$ km.

Since the eigenfunctions $\psi_n(\hat{\mathbf{r}})$ form an orthogonal functional basis, the covariance function can be expanded into an EOF form

$$\rho(\hat{\mathbf{r}}, \hat{\mathbf{r}}') = \sum_{n=1}^{\infty} \lambda_n \psi_n(\hat{\mathbf{r}}) \psi_n(\hat{\mathbf{r}}'). \tag{5.7}$$

The EOF representations of $\rho(\hat{\mathbf{r}}, \hat{\mathbf{r}}_i)$ and $\bar{\rho}_i$ are respectively

$$\rho(\hat{\mathbf{r}}, \hat{\mathbf{r}}_i) = \sum_{n=1}^{\infty} \lambda_n \psi_n(\hat{\mathbf{r}}) \psi_n(\hat{\mathbf{r}}_i) \tag{5.8}$$

and

$$\bar{\rho}_i = \sum_{n=1}^{\infty} \lambda_n \psi_n(\hat{\mathbf{r}}_i) \bar{\psi}_n, \tag{5.9}$$

where $\bar{\psi}_n$ is defined as

$$\bar{\psi}_n = \frac{1}{A} \int_{\Omega} \psi_n(\hat{\mathbf{r}})\, d\Omega \tag{5.10}$$

which is the average of the eigenfunction $\psi_n(\hat{\mathbf{r}})$. In practice one has to compute an approximate value of this by numerical integration:

$$\bar{\psi}_n \approx \frac{1}{A} \sum_{j \in \mathcal{OI}} \psi_n(\hat{\mathbf{r}}_j) A_j \approx \sum_{j \in \mathcal{OI}} \hat{v}_j^{(n)} \frac{\sqrt{A_j}}{A}. \tag{5.11}$$

In summary, to compute $\bar{\rho}_i$ we:

[1] Compute the covariance matrix $[\rho_{ij}^{OI}]$ according to (4.13) (excluding the term $\langle E_i^2 \rangle \delta_{ij}$) and the modified covariance matrix $\hat{\rho}_{ij}$ according to (5.3).

[2] Solve the eigenvalue problem for the modified covariance matrix $\hat{\rho}_{ij}$ to obtain eigenvalues $\hat{\lambda}_k$ and normalized eigenfunctions $\hat{v}_i^{(k)}$.

[3] Use (5.11) to compute $\bar{\psi}_n$ and (5.4) to compute $\psi_n(\hat{\mathbf{r}}_i)$, and finally compute $\bar{\rho}_i$ by

$$\bar{\rho}_i \approx \sum_{n=1}^{M_c} \hat{\lambda}_n \psi_n(\hat{\mathbf{r}}_i) \bar{\psi}_n. \tag{5.12}$$

The quantities $\hat{\rho}_{ij}$ and $\bar{\rho}_i$ will be used in eq. (4.11) which, together with eq. (4.12), determines the optimal weights $w_1, \cdots, w_N$ for averaging. The eigenvalues $\hat{\lambda}_n$, eigenfunctions $\psi_n(\hat{\mathbf{r}}_i)$, and their averages $\bar{\psi}_n$ will be used to calculate the total sampling error given by eq. (5.13) below. The sum in eq. (5.12) above and eq. (5.13) below for n in practice runs through a relatively small number of modes M_c, say 12 for OI+GHCN blended data, since the higher modes are contaminated by noise and the inclusion of these modes may increase error. As discussed by Kagan (1979) it is important to avoid adding more detail to the covariance function than can be justified by the amount of data available to compute them. However, this problem is lessened by using EOF and the MSE formula (5.13) below since each mode is scaled by its eigenvalue. This forces the first few, most important modes, to dominate. In practical

computations, one may choose the cut-off mode number M_c according to the criterion of 80 to 95% variance explained (by the first M_c modes).

By (4.9) and (5.7), the final expression of MSE is obtained in terms of M_c EOF modes

$$\begin{aligned}
\epsilon^2 &= \frac{1}{A^2}\int_\Omega d\Omega \int_\Omega d\Omega' \rho(\hat{\mathbf{r}},\hat{\mathbf{r}}') - 2\sum_{j\in\mathcal{N}} w_j\bar{\rho}_j + \sum_{i,j\in\mathcal{N}} w_i w_j\ \rho_{ij} + \sum_{i\in\mathcal{N}} w_i^2\langle E_i^2\rangle \\
&= \frac{1}{A^2}\int_\Omega d\Omega \int_\Omega' d\Omega' \sum_{m=1}^{\infty}\lambda_n\psi_n(\hat{r})\psi_n(\hat{r}') - 2\sum_{j\in\mathcal{N}} w_j\frac{1}{A}\int_\Omega \rho(\hat{r},\hat{r}_j')d\Omega \\
&+ \sum_{ij\in\mathcal{N}} w_i w_j \sum_{n=1}^{\infty}\lambda_n\psi_n(r_i)\psi_n(r_j) + \sum_{j\in\mathcal{N}} w_i^2\langle E_i^2\rangle \\
&= \sum_{n=1}^{\infty}\lambda_n\bar{\psi}_n^2 - 2\sum_{n=1}^{\infty}\lambda_n\sum_{j\in\mathcal{N}} w_j\psi_n(\hat{r}_j)\bar{\psi}_n + \sum_{n=1}^{\infty}\lambda_n(\sum_{j\in\mathcal{N}} w_j\psi_n(\hat{r}_j))^2 + \sum_{i\in\mathcal{N}} w_i\langle E_i^2\rangle \\
&\approx \sum_{n=1}^{M_c}\hat{\lambda}_n\left[\bar{\psi}_n - \sum_{j\in\mathcal{N}} w_j\psi_n(\hat{\mathbf{r}}_j)\right]^2 + \sum_{i\in\mathcal{N}} w_i^2\langle E_i^2\rangle. \qquad (5.13)
\end{aligned}$$

Since this formula includes the EOF patterns, if the observations are along the node lines of an EOF (where the EOF is equal to zero) or in the fine spatial structure area of an EOF, the sampling error is large for the corresponding mode. Thus this formula is also useful for future observation network design.

The sampling error formula (5.13) implies that the MSE is linearly proportional to the eigenvalues and is in a square relationship with the numerical integration errors of the eigenfunctions. This is the mathematical basis of many researchers' opinion that to estimate the MSE of an OA it is crucial to obtain highly accurate eigenvalues, and the exact shapes of the eigenfunctions do not matter as much. Therefore it is desirable to compute the eigenvalues as accurately as possible.

Chapter 6

Optimal averaging results

This section describes our computational results on eigenvalues, eigenfunctions, OA, its MSE and optimal weights. In the discussion of Chapter 5, the eigenvalue was believed to be the most important component in OA. The influence of eigenvalues on MSE is larger than that of eigenfunctions on MSE (5.13). Section 6.1 will give a detailed discussion of eigenvalues. An accurate eigenfunction is important in the optimal weights calculation, and the optimal weights will influence the result of OA. The optimal weights are important because the purpose of OA is to obtain an accurate annual average temperature. Section 6.2 will pay attention to the eigenfunctions. The OA result will be given in Section 6.3. Sections 6.4 and 6.5 will discuss errors and weights, respectively. Now we begin with the eigenvalues and their validation.

6.1 Eigenvalues and their validation

The area-weighted total variance $\langle \int_\Omega T^2(\hat{\mathbf{r}})\ d\Omega \rangle$ can be used as verification for the eigenvalue computation and as a reference quantity to determine the number of modes

used in sampling error calculations. Table 1 shows the eigenvalues in January. From these eigenvalues, one can calculate the total variance in January. This variance $\langle \int_\Omega T^2(\hat{\mathbf{r}})\, d\Omega \rangle$ can be computed directly from numerical integration

$$\langle \int_\Omega T^2(\hat{\mathbf{r}})\, d\Omega \rangle \approx \frac{1}{26} \sum_{\gamma=1}^{26} \sum_{j \in \mathcal{OI}} T^2(\hat{\mathbf{r}}_j, \gamma) \cos\phi_j \, (5\pi/180)^2 = 22.425 \;\; [^\circ\mathrm{C}]^2 R^2 \qquad (6.1)$$

where γ indicates the time variable in the unit of month, ϕ_j is the latitude of the grid point $\hat{\mathbf{r}}_j$, and $\cos\phi_j \, (5\pi/180)^2$ is the area of the jth $5^\circ \times 5^\circ$ grid box.

Because of the statistical similarity of the eigenvalues of the two datasets, only the eigenvalues resulting from the OI+GHCN blended data are discussed. From Table 1, one finds that the first 12 modes (less than half) can explain 86.85% of the total variance.

The summation of the first 12 of these eigenvalues is

$$\sum_{n=1}^{12} \hat{\lambda}_n = 19.52.$$

Because of

$$\langle T^2(\hat{\mathbf{r}}) \rangle = \sum_{n=1}^{\infty} \lambda_n \psi_n^2(\hat{\mathbf{r}})$$

and the normalization condition for $\psi_n^2(\hat{\mathbf{r}})$, one has

$$A\langle T_{(m)}^2 \rangle = \langle \int_\Omega T^2(\hat{\mathbf{r}})\, d\Omega \rangle = \sum_{n=1}^{\infty} \lambda_n \approx \sum_{n=1}^{12} \lambda_n = 19.52 \;\; [^\circ\mathrm{C}]^2 R^2,$$

where R is the radius of earth and $\langle T_{(m)}^2 \rangle$ is the mean value of the SST variance. The total variance 19.52 $[^\circ\mathrm{C}]^2 R^2$ computed from the summation of the first 12 eigenvalues is in good agreement with 22.425 $[^\circ\mathrm{C}]^2 R^2$ computed from the direct numerical integration shown in eq. (6.1). Note that this is an independent verification for the eigenvalue computation. The area of Ω is $A = 4\pi\ \mathrm{R}^2$. Hence

$$\langle T_{(m)}^2 \rangle \approx \frac{19.52}{4\pi} = 1.553 \;\; [^\circ\mathrm{C}]^2.$$

The standard deviation is

$$\sqrt{\langle T^2_{(m)} \rangle} = 1.246[^\circ\mathrm{C}]. \tag{6.2}$$

The first 12 extrapolated eigenvalues from the 5° ×5° grid explains 86.85 % (= 19.52 ÷ 22.425) of the total variance. Considering that the high modes contain much noise, it was decided to take the first 12 modes in our OA and sampling error estimation, i.e., $M_c = 12$ for OI+GHCN blended data in eqs. (5.12) and (5.13). Similarly, $M_c = 20$ was used for weaver data in the rest of this paper. This value also explains 85% of the total variance of the Weave data.

Another direct method to select the modes is to use the eigenvalue curve. Fig. 6 shows the eigenvalues for each mode in January. Examining Fig. 6, one can conclude that the contribution after mode 12 is very small compared to the first 12 modes. They can be considered as white noise after mode 12.

Now, the eigenvalues vary with the month. Comparing the eigenvalues in January (Table 2) and July (Table 3), the first mode in January explains 20.5% of the total variance while the first mode in July explains 28.1% of the total variance. But the summation of the first 12 modes explains a similar percentage of the total (86.85% in January, 89.17% in July). The difference comes from the length of the sample for EOF analysis: there are 26 modes in total in January EOF analysis and 24 modes in total in July EOF analysis. So the percentage of 12 modes in July is a little bit higher than the one in January.

6.2 Eigenfunctions

Now we turn to the eigenfunction discussion. Eigenfunctions of a covariance function often reflect the field's physical patterns. Through analyzing the eigenfunctions, one

can infer the surface temperature distribution characteristics, such as ENSO. Fig. 7a shows the first 6 eigenfunctions in January for OI+GHCN blended data. One can find that each eigenfunction reflects the land and sea distribution well in January. The small scale structures in each mode arise from the inconsistency of OI, GHCN blended data and Reanalysis data. The second characteristic is the ENSO phenomenon. One can see that there are 5 modes containing ENSO phenomena in January (all except the second mode). This is the result of the fact that the ENSO mode was the main mode in the tropical ocean in January during the 1980's and 1990's. This is consistent with the frequent occurrence of strong *El Niño* in the 1980's and 1990's. This characteristic is also reflected in July (Fig. 7b). Looking at Fig. 7b, one can find the difference in eigenfunctions between January and July. In July, not only is the land-sea difference not as strong as in winter (because the winter system in Northern Hemisphere is very strong, such as the winter monsoon in Asia), but also the ENSO modes are weaker than the ones in January. The ENSO modes only appear in modes 2, 4, 5 and 6. Comparing the January and July eigenfunctions (Fig. 7a and Fig. 7b), one forms the impression that the strongest ENSO signal mainly emerges in Northern Hemispheric winter. In January, the *Nino* 3, which is bounded by 90°W-150°W and 5°S-5°N, is strong for first six modes. But in July, the *Nino* 1+2 is the main part. This is consistent with the ENSO characteristic that the ENSO in summer (Northern Hemisphere) is relatively weaker than in winter (Northern Hemisphere).

One result from Section 6.1 is that the eigenvalues do not change significantly with seasons. This is true for different datasets, such as weaver of Jones and Reanalysis data for EOF analysis. What about the eigenfunctions? Their stability is investigated by comparing eigenfunctions belonging to different datasets. Fig. 8a and 8b show the first six eigenfunctions for the weaver data in January and July, respectively. The strength

of the ENSO mode for the weaver data is weaker than that for the OI+GHCN blended data. The reason is that the 1997-1998 El Niño was the strongest in this century and the ENSO in the 1980's and 1990's were stronger than the earlier period. It should be recalled that the OI+GHCN blended data use the data of 1980's and 1990's. When using these data for EOF, the eigenfunctions reflect the stronger ENSO mode. But the weaver data has a longer sample. This means that there is a smaller contribution to the total variance for the first six modes compared to the OI+GHCN blended data, because the ratio of the first 6 modes for weaver data is only 6/50 whereas the ratio for OI+GHCN blended data is 6/26. So the stronger ENSO mode in the first six modes of the OI+GHCN blended data is perceptible. There are also five ENSO modes in the first six modes shown in the weaver dataset in January. Because of the smoothing inherent in the weaver data, the small scale structures do not appear in the EOF of this dataset. Mode 6 still clearly contains the ENSO component not only in January but also in July. However, their structures are different. The SST anomaly over *Nino* 3 is obvious in January, but that over *Nino* 1+2 is strong in July. In conclusion, the eigenfunctions change very little with season even though the percentage variance of each mode is different. The weak magnitude can be compensated by increasing the mode and keeping the mode percentage the same. This is important for OA and data reconstruction, which will be discussed further in detail in Chapter 7.

6.3 Optimal averaging results

In this section, the previous discussion will be verified. First, let us discuss how many modes should be used. Section 6.1 demonstrated that eigenvalues do not change with respect to season, and it was determined that one should use 12 modes to do the OA

for the OI+GHCN blended data and 20 modes for the weaver data. A test of the mode selection will be done through calculations in this section. Two experiments are designed in this test to use the OA method twice for the same EOF analysis. The first experiment uses all 26 modes in the optimal weight calculation, whereas the second uses only 12 modes. Fig. 9 gives the comparison. One may easily see that the difference is very small (0.03°C). This means that using 12 modes is sufficient to perform the OA, with the modes after 12 regarded as white noise. When all modes are considered, this white noise is included and consequently weakens the optimal result. From the third panel in Fig. 9, the white noise influence can be observed. In this study, it ranges from -0.03°C to 0.03°C. This characteristic is also reflected in the weaver data for EOF analysis. 20 out of 50 modes are taken for OA using weaver data in the EOF analysis.

Now we consider the OA results from the OI+GHCN blended data for EOF analysis. Fig. 10 shows the comparison between OA and the Jones area-weighted average. The first panel shows the annual average comparison and its 10-year running average is given in the second panel. One can see that the OA and area-weighted average reflect the temperature trend synchronously. This is because of the constraint on the weight calculation for the two methods, i.e., $\sum w_i = 1$. Two periods of warming are clearly shown in panel two, namely 1920 - 1944 and 1978 - 1998. The difference between the two series is shown in the third panel. The solid line represents the annual difference, and the dashed line represents the 10-year running average for the difference. This figure shows that OA before 1980 is mostly higher than the Jones area-weighted average, while OA is lower than the Jones area-weighted average during the 1980's and 1990's. The difference is between -0.10°C and 0.15°C. This means the global warming in OA is not as strong as that in area-weighted average. Why did the area-weighted average

obtain so strong a global warming before?

Jones and his group made a series of studies of global surface temperature. In his 1988 and 1994 papers, in which he studied the land surface air temperature, the annual average result in the 1994 paper is higher than in the 1988 paper in the early period. The differences are mainly due to the introduction of a few extra Australian stations in Jones' 1994 work, which resulted in warmer late 19th century temperatures relative to the earlier time-series (IPCC 1995). Returning to the OA result (Fig. 10), the warmth in the late 19th century is very clear. The OA advantage is obvious here because the OA considers the influence of the structure. For example, if there is one grid point in Australia and the pattern there is large enough, the contribution for the anomaly temperature from this one point can represent the total temperature variance of Australia through a change in the weight distribution. Thus, through redistributing the weight, one can get a more accurate annual average than the simple area-weighted average, especially for the earlier period. In the IPCC report (1995), the improved data is typically up to 0.05°C warmer before 1900. Our result shows this difference will reach about 0.1°C and the average is about 0.03°C before 1900. Yet the difference between the area-weighted average and OA is small in the period of EOF (1982-1998). There are two possible reasons for this small difference. The first was proposed by Rayner (1999 personal communication), who believed the EOF influence is a crucial element in the OA method. That is, in her opinion, the OA result mainly depends on the EOF analysis. A second reason for the difference in question considers data quality. It basically suggests that the large amount of satellite data introduced after 1980 causes the difference between OA and area-weighted average to be small. In our later discussion, we will find that this is a consequence of the improvement of data quality. If this small difference comes from such an improvement of the data, one can

say the eigenfunctions are stable. This leads to the problem of whether we now can use the eigenfunctions since the year 1982 to reflect the patterns of the earlier year.

In Section 6.2, we touched lightly on the question of eigenfunction stability. It seems that the eigenfunctions are stable even though the mode amplitudes vary greatly. But we need more to prove this stability, and another experiment was designed for this comparison. This time, the weaver data were used for the EOF analysis, with the same method to do the OA. Because there is a long period for the EOF analysis (50 years), the hypothesis is that a small difference between the OA and the area-weighted average for the EOF period (1949-1998) will be obtained. Fig. 11 is the OA temperature time series using the weaver data for the EOF analysis. It is surprising to see that the difference between the OA and the area-weighted average is not small during the period of EOF analysis (1949-1998), although it is small in the 1980's and 1990's. This supports the above hypothesis that the small difference between the OA and the area-weighted average in the 1980's and 1990's is due to the improvement of the data quality. In the 1980's and 1990's, especially for the ocean data, a large body of satellite data were introduced and this improved the data quality. Consequently, the difference between the OA and the area-weighted average becomes small in this period. A further point is that in spité of the fact that the satellite data at high latitudes are poor because of snow cover, the OA result will not be influenced because there are no data coming from the same area. Yet one should emphasize that the OA method does not have this advantage over the area-weighted average when we have high quality data. Similar to the OA result for the OI+GHCN blended data, the late 19th century warmth is also obvious in this OA result, which means the global warming is not as strong as the area-weighted average result. The global warming over the period 1856-1998 is about 0.05°C less than previously inferred. Another difference between the

OA result using the weaver data and the one using the OI+GHCN blended data is that the former result is not always larger than the area-weighted average before 1980. Also, the amplitude of the difference is less than that using the OI+GHCN blended data in the EOF analysis, about -0.12°C to 0.12°C. Even though the EOF analysis is stable, the longer dataset should be used to do the EOF analysis, since the longer the data sample, the more stable is the EOF analysis.

From the above discussion, two conclusions can be drawn in this section. First, the difference between the OA and the area-weighted average will become small when the data quality is improved. Second, the EOF analysis is very important and thus, if it is possible, one should use a long dataset to do the EOF analysis.

6.4 Error estimation

Error estimation is an important part in determining the global temperature anomaly. Through estimating the error, one can learn the global average accuracy. Errors provide us useful information for improving our method of estimating the global temperature anomaly.

In this study, the error is composed of two parts from eq. (5.13): sampling error and observational error. Sampling error comes from the spatial gaps in the data, whereas observational error accounts for the error in the $5° \times 5°$ box. Eq. (5.13) is used to estimate the sampling error and the observational error separately. Fig. 12 shows the error changing with time, and the total error decreasing with time. The error becomes larger during the two World Wars because of the decreasing sample size. The total error amplitude is about 0.05°C during the early period and 0.01°C in recent times. From Fig. 12, one obvious property is that the sampling error is less than

the observational error not only in January but also in July, because optimization is used for minimizing the sampling error. Another property is that not only the observational error but also the sampling error in January are larger than those in July. The sampling error using the weaver data for the EOF analysis is a little bit higher than that using the OI+GHCN blended data, peaking at 0.03°C during the earlier period (Fig. 13). Comparing the modes (20/50) used in the weaver data, this error is reasonable, because the OI+GHCN blended data use half of the modes in OA. Fig. 14 shows the area-weighted error. Unlike the OA error, the error for the area-weighted average does not decrease stably with time. Its amplitude is about 0.06°C to 0.18°C. This is the same as the amplitude of the difference between the OA and the area-weighted average. So the difference between the OA and the area-weighted average is due to the sampling error. Thus the smaller the sampling error, the more accurate the annual average will be. Fig. 14 also shows that the amplitudes of the sampling error and the observational error switch in the area-weighted average, that is, the sampling error is larger than the observational error. Because the area-weighted calculation does not perform an optimization for sampling error, the sampling error becomes very large compared to the optimal method. Another problem is that the most of the sampling error for the area-weighted average in January is less than the most of the sampling error in July. This is also switched compared to the OA error result: from Fig. 12, the OA error is small in July and large in January.

6.5 Weight distribution

Because the OA redistributes weights, the optimal weights play an important role in the OA result. In this section, the weight distribution will be discussed and the

evidence for weight influence on the global average will be shown.

Figs. 15a and 15b show the weight distribution in 1900 and 1983, respectively. As opposed to the area weight, the optimal weight is inhomogeneous. Comparing Figs. 15a and 15b, one can see that the weight distribution is similar in both years. This consistency shows the influence of covariance structure. If there is an accurate structure (EOF), there will be accurate weights, and the accurate weights will lead to an accurate global average temperature. For example, comparing Fig. 4a and Fig. 15a, one common characteristic for both periods is the positive anomaly temperature but negative optimal weights in the United States. This means that the anomaly temperature in the United States has a negative contribution to global warming. This is a major difference with the area-weighted average because the area weight is always positive. Thus, for the area-weighted average, the positive anomaly is always a positive contribution to the global average anomaly. Consequently, the inhomogeneous weight distribution in the OA method shows the OA's advantage. That is, it provides a better reflection of the anomaly temperature structure than the area-weighted average. The same phenomenon occurs when using the weaver data to calculate EOF (Fig. 16). Because the weaver data has a longer history, the strength of each mode is less than OI+GHCN blended data. Consequently, the weights are weaker than when using the OI+GHCN blended data for the EOF. This means that the contribution of negative weights is not as strong as the weights of the OI+GHCN blended data; however, the negative contribution is still obvious, especially in the southern United States.

In order to clearly show the weight contribution to the global temperature anomaly, two more figures are presented on this topic. Fig. 17a and Fig. 17b show the product of the weight and the temperature anomaly for the two years. The contribution of single box to the global anomaly is clearly shown in these two figures. It depends

not only on the temperature anomaly but also on the weight. Even a local negative anomaly can contribute to the positive anomaly for the global average. Comparing the anomaly temperature (Fig. 4a) and Fig. 17a, one can see that even though there are positive anomalies in the United States in 1900, the optimal result is not positive in this region. A negatively weighted temperature appears in the middle of the United States. One can also see negative weighted temperature along the East and West coasts. On the other hand, in Russia, the negative anomaly temperature in Fig. 4a corresponds to the negative anomaly in Fig. 17a. In 1983, similar results occurred. In the middle of the United States, negative weighted temperatures are obvious, as well as along the East and West coasts. Once again the weighted temperature in Russia is positive, corresponding to the positive anomaly temperature (Fig. 4b). One may say the contribution of the anomaly temperature from Russia to the global anomaly temperature corresponds to the local anomaly temperature variation, but the contribution of the anomaly temperature from the United States to the global anomaly temperature is opposed to the local anomaly temperature variation in 1900 and 1983.

Now that the global anomaly average has been studied, another question is whether the contribution of each latitude is the same. Zonal mean anomalies of combined land-surface air temperature (from Jones 1994) and SST (from Parker et al. 1995) confirm that the recent warmth is greatest in mid-latitudes of the Northern Hemisphere. This can be compared with the warm period of the mid-20th century, where the greatest warmth was in the high latitudes of the Northern Hemisphere. The recent warm period also exhibits higher temperatures in the southern Hemisphere (IPCC, 1995).

In order to check the consistency and differences between our results and those of Jones and Parker, we divide the globe into 6 bands. Each band spans 30 degree of latitude. The results of averaging over each band are shown in Fig. 18. In this figure,

one can see that the global warming mainly comes from 0-60°N, especially in the middle latitude of the Northern Hemisphere (0-360, 30°N-60°N). Most contributions to the first period warming are in the band of 0°N-30°N. For the second period of warming (1978-1998), most contributions come from 30°N-60°N. The contributions from the other latitude bands to global warming are not obvious.

The Pacific decadal time-scale variations have been linked to recent changes in the frequency and intensity of El Niño versus La Niña events, and it has been hypothesized that the decadal variation has its origin in the tropics (Trenberth and Hurrell, 1994). Observational studies by Kawamura (1994), and Lau and Nath (1994) have shown that the decadal variation in the extratropics of the Pacific is closely tied to tropical sea surface temperatures in the Pacific and Indian Oceans. Several aspects of the decadal-scale fluctuations beginning around 1976 have been simulated with atmospheric models (Kitoh 1991, Chen et al. 1992, Miller et al.1994, Kawamura et al. 1995). These studies also suggest that the temperature changes over the North Pacific are substantially controlled by the anomalous SST forcing from the tropics.

Fig. 18 also shows this decadal time-scale anomaly temperature variation. The long-term variation in 30°S-30°N temperature is more obvious than in mid-high latitudes. In the mid-high latitudes of the Northern Hemisphere, not only does the variance of the anomaly temperature become larger but the time scale also becomes longer. It revealed regions with strongly enhanced variability. The El Niño events after the 1980's enhanced the variance in the eastern tropical Pacific from 1978 to 1998 relative to the variance from 1950 to 1980.

Chapter 7

Conclusion and Discussion

The OA method presented in this work makes use of the EOF derived from the recently available OI+GHCN blended data and weaver data. Our results include not only the spatial average of the historical data from 1856 to 1998 but also the standard error. The standard error is measured by RMSE and is computed from the theoretical formula eq. (5.13). It is concluded that the OA result is reliable and accurate, and it takes into account the spatial inhomogeneity of the data. The advantages and disadvantages of OA will be discussed in Section 7.1 and 7.2. In Section 7.3, conclusions will be given.

7.1 Stability of EOF

The stability of EOF was discussed in Sections 6.1 and 6.3. Section 6.1 mainly discussed eigenvalue stability and Section 6.3 discussed eigenfunction stability. However, other aspects of EOF stability need further discussion.

Gray (1981) suggests that similar patterns should occur in the same order if the ratio of time to space is high (greater than 10). In Wigley's study (1984), the same

patterns occur in the same order down to mode 4, and Tabony (1981) obtained similar results. A question is that the above authors used a long time series to study a small area, hence, they have a high ratio of time to space. But in our dataset, the short period of the data for EOF analysis is obvious, while the spatial dimensions are very large. The spatial size is 72×36, which is much more than the temporal length (26 for the OI+GHCN blended data and 50 for the weaver data). How to deal with this kind of problem? Smith et al. (1996) suggested a method that divides the globe into several areas and performs the anomaly temperature reconstruction in each area. But there is a problem in using this method. If the globe is divided into several bands according to Smith et al. (1996), the stationarity for each band is better than for the global data because of the higher ratio of time to space. This means similar patterns may be obtained even for higher modes during different periods. However, such a procedure introduces another problem. For example, some regions do not have enough data or had no data during the early period. As shown in Fig. 4a, there were less data in the Pacific ocean and at high latitudes at the beginning of this century than at present. Such a sparse dataset will influence the accuracy of OA, or prevent OA analysis at all where no data are available.

In order to solve the problem of the lack of data in the early time, global data are used to do the OA. This introduces the problem of nonstationarity. Is there a way to minimize this nonstationarity? The following experiment will give the answer. Because the Reanalysis data have quite a long history, starting from 1949, they will be used to do the experiment. First, we divide the Reanalysis into two datasets, one from 1949 to 1973, and the other from 1974 to 1998. Each covers a period of 25 years. Next, we perform the EOF analysis for the entire Reanalysis dataset from 1949 to 1998. In order to explain the eigenfunctions clearly, the first 12 modes are shown for

each dataset. The differences are obvious for the two datasets. Figs. 19a and 19b show the eigenfunctions for 1949-1973 in January. The modes reflect the land-sea distribution well. The 5^{th}, 8^{th} and 9^{th} modes reflect the ENSO component. One can see that the ENSO mode is weak during the first 25-year dataset, but strong for the second 25-year dataset. Figs. 20a and 20b show the first 12 modes for the second 25 years. Besides the land-sea distribution, the ENSO pattern is evident in modes 2, 5, 6, 7, 8, 9, 11 and 12. Because the eigenvalues decrease with modes, the ENSO modes for the second 25 years are stronger than for the first 25 years. This is consistent with the observation that the second 25-year period contains stronger ENSO events, with more frequent occurrence during 1980's and 1990's.

Fig. 21 shows the eigenfunctions for the entire period, 1949-1998. The ENSO pattern is reflected in modes 2, 4, 8, 9, 10 and 12. The whole dataset exhibits some common characteristics in both the first 25 years and the second 25 years. Using these three datasets for EOF analysis, one gets three OA results. There are obvious differences between the first 25-year time series and the second 25-year time series shown in Fig. 22. Compared to this large OA difference, the difference between the first 25-year and the 50-year EOF analyse is small (Fig. 23). Because the 50-year EOF analysis contains all the characteristics of the first and second 25-year EOF analyses, while the longer EOF analysis is more stable, in further analysis, one should use longer series to do the EOF analysis. The stability discussed here is different from the stability discussed in Section 6.3. There, the differences between EOF period and non-EOF period OA are discussed. But in this section, a single dataset is divided into two parts. If it is possible, the longer period dataset should be used.

7.2 Comparison between optimal averaging and spatial interpolation

Smith et al. (1996) discussed the advantages and disadvantages of OA versus spatial interpolation. He pointed out that traditional spatial and temporal interpolation into data void regions can give larger errors, whereas EOF reconstruction gives a more spatially coherent field of SST anomalies. The disadvantage of the traditional method can be ameliorated by the reconstruction method, especially in the tropical Pacific, where anomalies can be large and persistent and where sampling is often sparse. In regions where in situ sampling is dense, the EOF based reconstruction does not have as clear an advantage over more traditional methods of forming gridded fields. In this paper, the OA method shows its advantage in the 19th century. It considers the spatial structure and minimizes the error optimally. However, the OA method has less advantage in determining recent global average temperature anomalies, especially for the last two decades, when high quality and dense data from satellites have been available.

7.3 Conclusions

From the above discussion, the following conclusions can be drawn:

(i) The longer dataset should be used for EOF analysis. The longer the EOF dataset, the more accurate the annual average.

(ii) An increase in the length of the data period for EOF analysis cannot reduce the difference between the OA and the area-weighted average. The reduction of this difference depends on data quality and density improvement.

(iii) The standard error of the OA is less than that of the area-weighted average. The amplitude of the difference between the OA and area-weighted average is similar to the area-weighted sampling error.

(iv) The optimal weight field is important for estimating global temperature variations. Unlike the area-weighted average, the optimal weight changes the contribution of each area to the global temperature variation. The positive anomaly is always positive for the area-weighted average, but this is not true for the OA. As discussed in Section 6.5, the positive anomaly may contribute a negative anomaly to the global average because of the opposing sign of anomaly temperatures and weights. The introduction of negative weights shows the advantage of OA, because it considers not only the data values but also their physical structure.

REFERENCES

Chen, L-X., M. Dong and Y.-N. Shao, 1992: The characteristics of interannual variations of the East Asia Monsoon. *J. Met. Soc. Japan* , **70**, 397-421.

Cook, E.R., K.R. Briffa, and P.D. Jones, 1994: Spatial regression methods in dendroclimatology: a review and comparison of two techniques. *Int. J. Climatol.*, **14**, 379-402.

Folland, CK, and D.E. Parker, 1995: Corrections of instrumental biases in historical sea surface temperature data, *Q. J. R. Meteo. Soc.*, **121**, 319-367.

Gray, B.M., 1981: On the stability of temperature eigenvector patterns. *Journal of Climatology.*, **1**, 273-281.

Hansen, J., and S. Lebedeff, 1987: Global trends of measured surface air temperature. *J. Geophys. Res.*, **92**, 13345-13372.

IPCC, 1995: Climate Change. The Science of Climate Change. *Cambridge.*, 132-227.

Jones, P.D., S.C.B. Raper, R.S. Bradley, H.F. Diaz, P.M. Kelly, and T.M.L. Wigley, 1986a: Northern Hemisphere surface air temperature variations, 1851-1984. *J. Climate Appl. Meteor.*, **25**, 161-179.

Jones, P.D., and T.M.L. Wigley, 1986b: Southern Hemisphere surface air temperature variations: 1851-1984.*J. Climate Appl. Meteor.*, **25**,1213-1230.

Jones, P.D. 1994: Hemispheric surface air temperature variation: a Reanalysis and an update to 1993. *J. Clim*, **7**, 1794-1802.

Jones, P.D., T.J. Osborn and K.R. Briffa 1997: Estimating sampling errors in large-scale temperature *J. Clim.*, **10**,2548-2568.

Jones, P.D., M. New, D.E. Parker, S. Martin and I.G. Rigor 1999: Surface air temperature and its changes over the past 150 years. *Rev. Geophys.* ,**37**, 173-199.

Kagan, R.L., 1979: *Averaging Meteorological Fields.* Gidrometeoizdat, 212 pp. (in Russian). [English translation is available from National Center for Environmental Modeling, National Weather Service, NOAA, Washington, DC 20233.]

Kaplan A., Y. Kushnir, M.A. Cane and M.B. Blumenthal, 1997: Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures. *J. Geophys. Res.*, **102**, 27835-27860.

Kawamura, R.,1994: A rotated EOF analysis of global sea surface temperature variability with interannual and interdecadal scales. *J. Phys. Oceanogr.*, **24**, 707-715.

Kawamura, R.,M. Sugi and N. Sato, 1995: Interdecadal and interannual variability in the northern extratropical circulation simulated with the JMA global model. *J. Clim.*

Kim, K. -Y., and G.R. North, 1993: EOF analysis of surface temperature field in a stochastic climate model. *J. Climate* **6**, 1681-1690.

Kim, K. -Y., G.R. North and S.S. Shen, 1996: Optimal estimation of spherical harmonic components from a sample with spatially nonuniform covariance statistics. *J. Climate* **9**, 635-645.

Kitoh, A., 1991: Interannual variations in an atmospheric GCM forced by the 1970-1989 SST. Part II: Low frequency variability of the wintertime Northern Hemisphere extratropics. *J. Met. Soc. Japan*, **69**, 271-291.

Lau, N.-C. and M.J. Nath, 1994: A modeling study of the relative roles of tropical and extratropical SST anomalies in the variability of the global atmosphere-ocean system. *J. Clim.*, **7**, 1184-1207.

Madden, R.A., D.J. Shea, G.W. Branstator, J.J. Tribbia, and R.O. Weber, 1993: The effects of imperfect spatial and temporal sampling on estimates of the global mean temperature: Experiments with model data. *J. Clim.*, **6**, 1057-1066.

Miller, A.J., D.R. Cayan, T.P. Barnett, N.E. Graham and J.M. Oberhuber, 1994: Interdecadal variability of the Pacific Ocean: Model response to observed heat flux and wind stress anomalies. *Clim. Dyn.*, **9**, 287-301.

Oort, A.H.,1978: Adequacy of the rawinsonde network for global circulation studies tested through numerical model output.*Mon. Wea. Rev.*, **106**, 174-195.

Parker, D.E., C.K. Folland, and M. Jackson, 1995: Marine surface temperature: observed variations and data requirements. *Climatic Change* **31**, 559-600.

Rayner N.A., 1999: Optimum Averaging: Next Steps. *Hardley Center for Climate Prediction and Research, Met. Office, U.K.*, Personal communication

Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. *J. Clim.*, **1**, 75-86.

Reynolds, R. W., 1993: Impact of Mount Pinatubo aerosols on satellite-derived Sea Surface Temperatures. *J. Clim.*, **6**, 768-774.

Reynolds, R. W. and D. C. Marsico, 1993: An improved real-time global sea surface temperature analysis. *J. Clim.*, **6**, 114-119.

Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analysis using optimum interpolation. *J. Clim.*, **7**, 929-948.

Shen, S. S., G. R. North, and K.-Y. Kim, 1994: Spectral approach to optimal estimation of the global average temperature. *J. Climate* **7**, 1999-2007.

Shen, S.S., T.M. Smith, C.F. Ropelewski, and R.E. Livezey, 1998: An optimal regional averaging method with error estimates and a test using tropical Pacific SST data *J. Clim.*, **11**, 2340-2350.

Smith, T.M., R.W. Reynolds, and C.F. Ropelewski, 1994: Optimal averaging of seasonal sea surface temperatures and associated confidence intervals (1860-1989). *J. Clim.*, **7**, 949-964.

Smith, T.M., R.W. Reynolds, R.E. Livezey, and D.C. Stokes, 1996: Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *J. Clim.*, **9**, 1403-1420.

Tabony, R.C., 1981: A Principal Component and Spectral Analysis of European Rainfall. *J. Climatol.*, **1**, 283.

Trenberth, K.E., and J.G. Olson, 1991: Representativeness of 63-station network for depicting climate changes. Greenhouse-Gas-Induced Climate Change: A Critical Appraisal of Simulations and Observations, M.E. Schlesinger, Ed., Elsevier, 615pp.

Trenberth, K.E., and J.W. Hurrell, 1994: Decadal atmosphere-ocean variations in the Pacific. *Clim. Dyn.*, **9**, 303-319.

Vinnikov, K. Ya., P. Ya. Groisman, and K.M. Lugina, 1990: Empirical data on contemporary global climate changes (temperature and precipitation). *J. Clim.*, **3**, 662-677.

Weber, R.O. and R. A. Madden, 1995: Optimal averaging for determination of global mean temperature: Experiments with model data. *J. Clim.*, **8**, 418-430.

Wigley, T.M.L., J.M. Lough and P.D. Jones 1984: Spatial patterns of precipitation in England and Wales and a revised, homogeneous England and Wales precipitation series. *J. Clim.*, **4**, 1-25.

Zwiers, F. W., and S. S. Shen, 1997: Errors in estimating spherical harmonic coefficients from partially sampled GCM output. *Clim. Dyn.*, **13**, 703-716.

Table 2 Modes 1 to 26 in January

Modes	Eigenvalues	Percentage of variance	Cumulative percent
1	4.60	20.47	20.47
2	3.24	14.41	34.88
3	2.34	10.43	45.31
4	1.82	8.12	53.42
5	1.40	6.24	59.66
6	1.24	5.51	65.17
7	1.18	5.25	70.42
8	.97	4.31	74.73
9	.90	4.00	78.72
10	.72	3.21	81.93
11	.60	2.66	84.59
12	.51	2.25	86.85
13	.44	1.97	88.82
14	.40	1.79	90.60
15	.37	1.64	92.24
16	.32	1.42	93.66
17	.28	1.26	94.92
18	.28	1.24	96.16
19	.22	.97	97.13
20	.14	.62	97.76
21	.11	.50	98.26
22	.10	.44	98.70
23	.09	.41	99.11
24	.08	.35	99.45
25	.06	.28	99.74
26	.06	.26	100.00

Table 3 Modes 1 to 24 in July

Modes	Eigenvalues	Percentage of variance	Cumulative percent
1	6.08	28.06	28.06
2	3.48	16.07	44.13
3	2.04	9.43	53.56
4	1.47	6.78	60.34
5	1.31	6.04	66.38
6	1.01	4.65	71.03
7	.94	4.34	75.37
8	.80	3.68	79.05
9	.64	2.95	82.00
10	.57	2.63	84.63
11	.50	2.28	86.92
12	.49	2.25	89.17
13	.45	2.07	91.24
14	.38	1.75	92.99
15	.33	1.51	94.50
16	.32	1.45	95.95
17	.28	1.31	97.26
18	.23	1.05	98.31
19	.11	0.51	98.83
20	.06	0.29	99.12
21	.06	0.28	99.40
22	.05	0.23	99.63
23	.04	0.20	99.82
24	.04	0.18	100.00

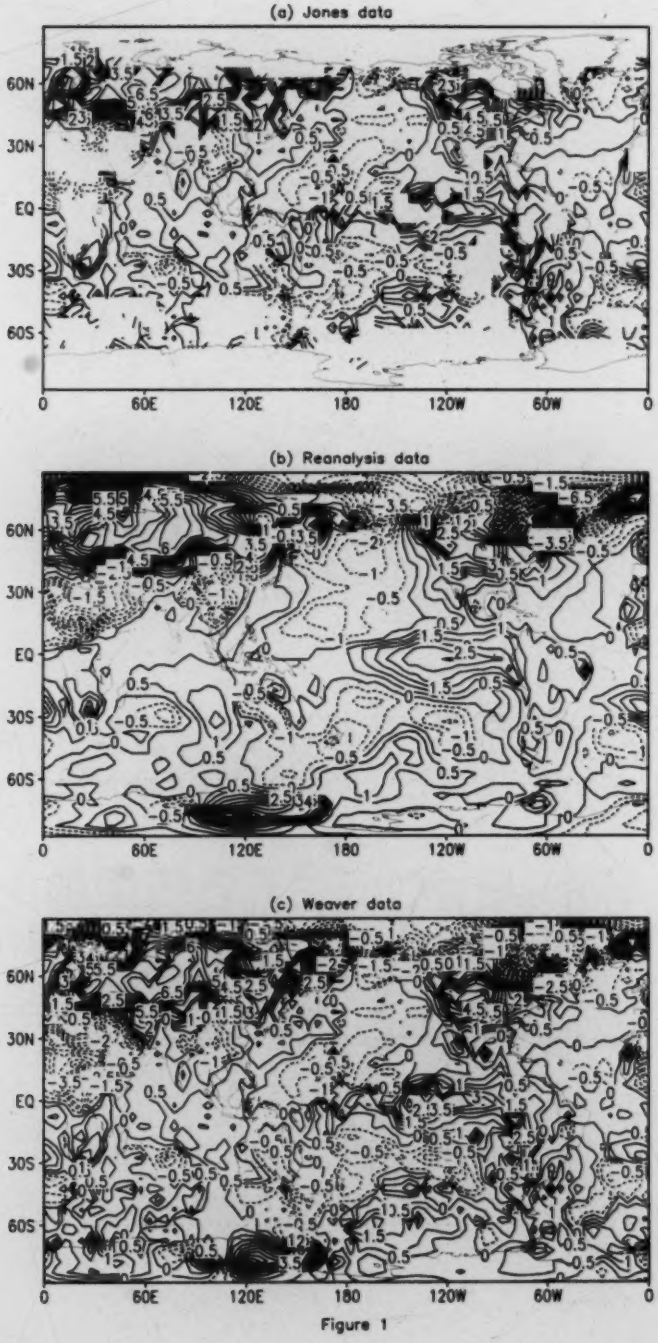

(a) Jones data
(b) Reanalysis data
(c) Weaver data
60N
30N
EQ
30S
60S
0
60E
120E
180
120W
60W

Figure 1

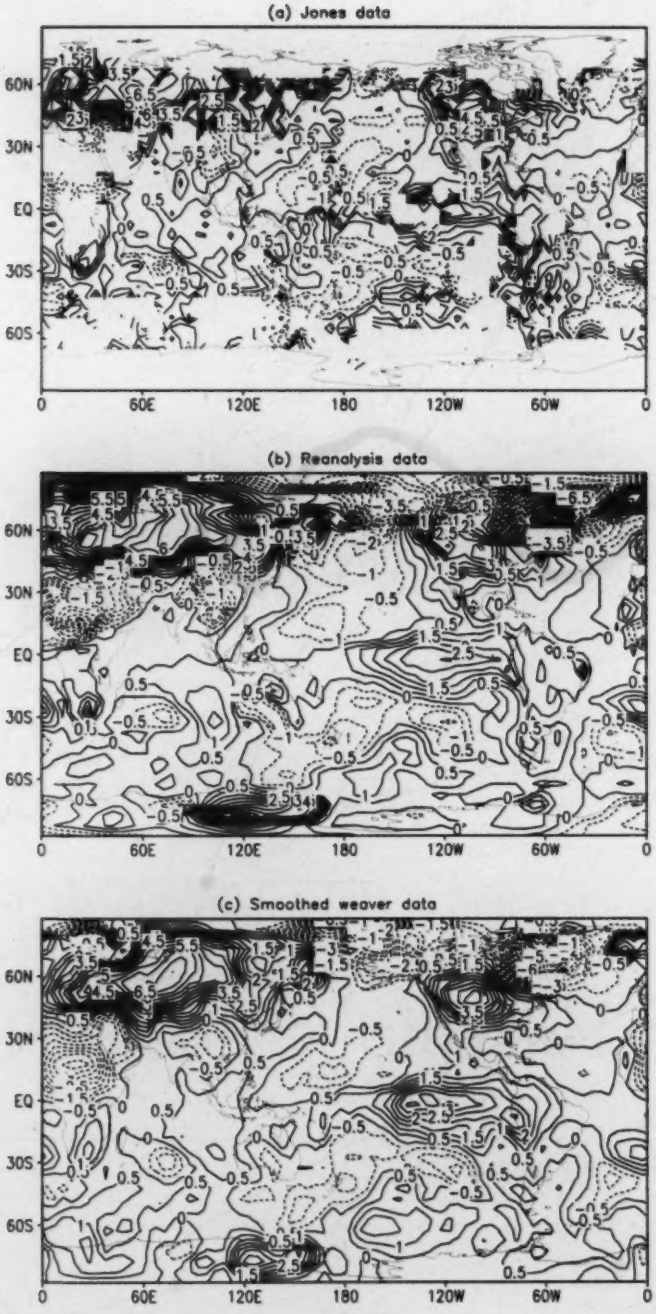

(a) Jones data
(b) Reanalysis data
(c) Smoothed weaver data
60N
30N
EQ
30S
60S
0
60E
120E
180
120W
60W

Figure 2

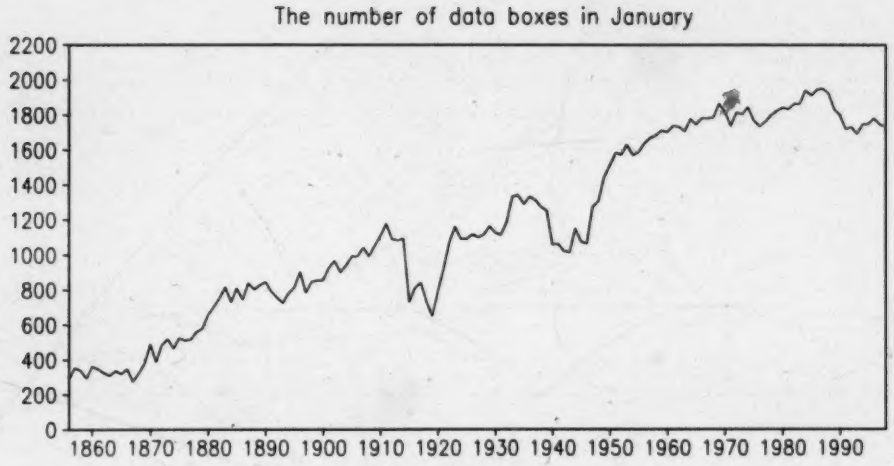

The number of data boxes in January
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990

Figure 3

The global anomaly temperature distribution in January 1900

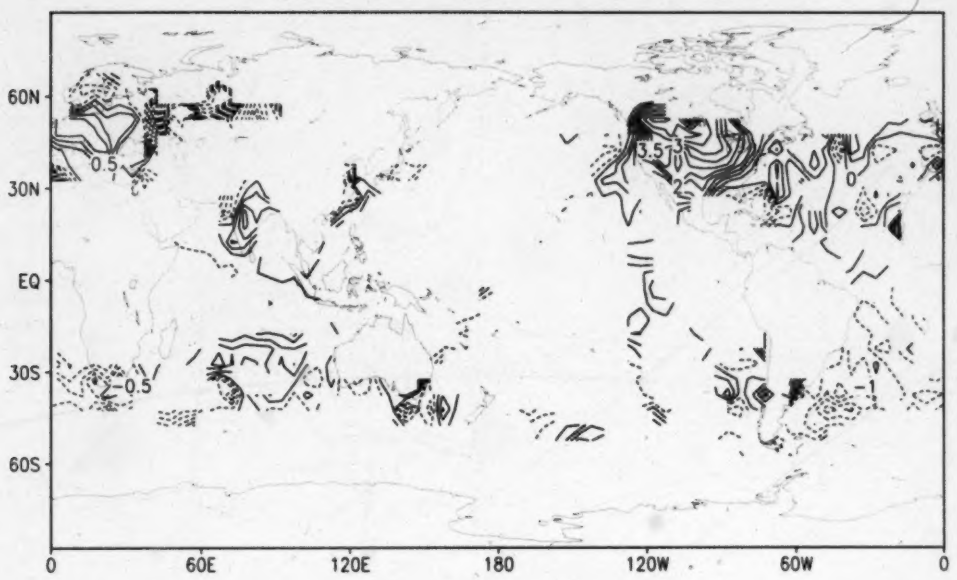


Figure 4a

The global anomaly temperature distribution in January 1983

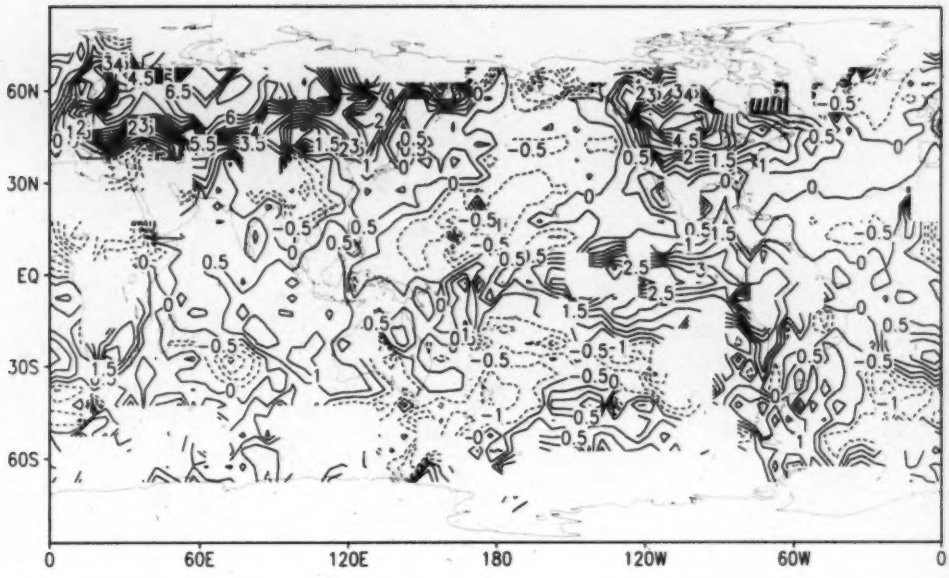


Figure 4b

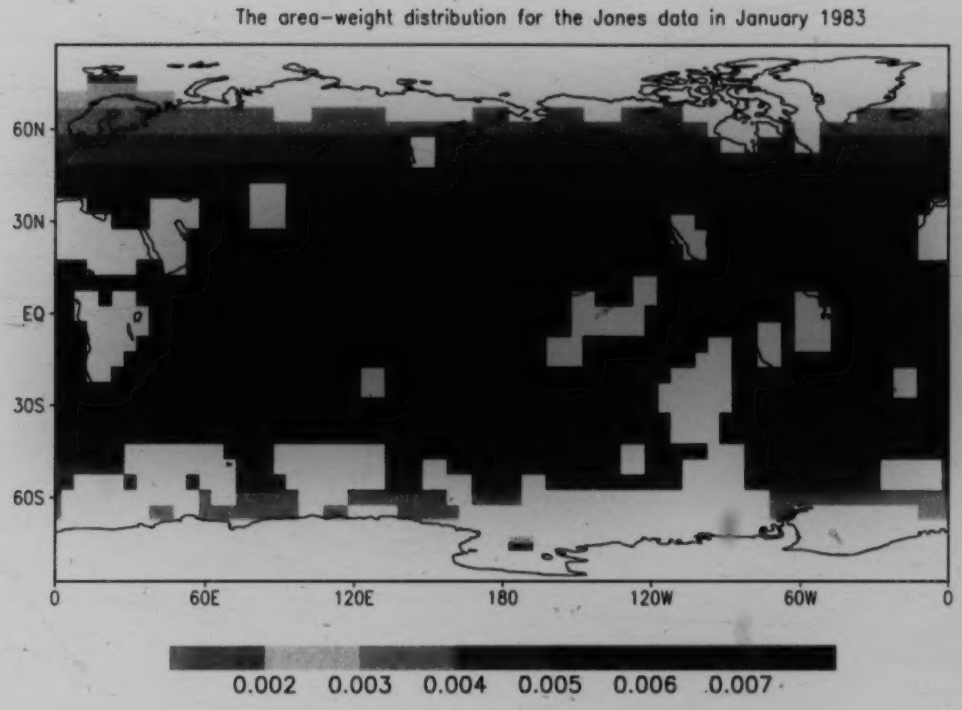

The area-weight distribution for the Jones data in January 1983
60N
30N
EQ
30S
60S
0
60E
120E
180
120W
60W
0
0.002
0.003
0.004
0.005
0.006
0.007

Figure 5

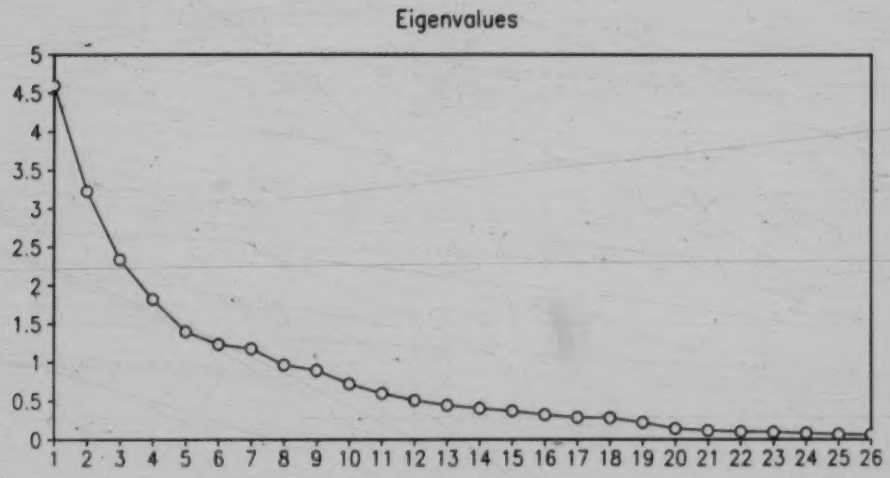

Eigenvalues
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

Figure 6

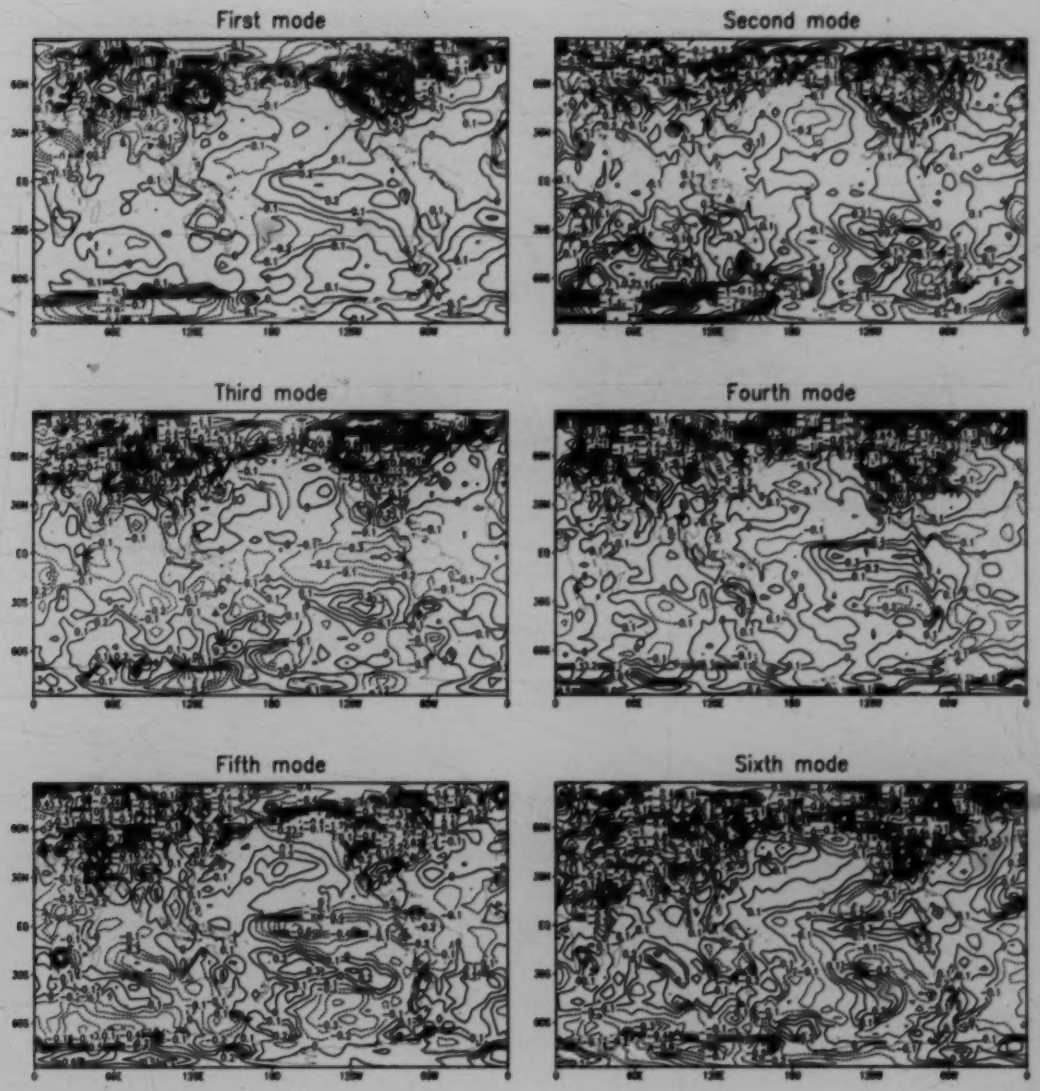

First mode
Second mode
Third mode
Fourth mode
Fifth mode
Sixth mode

Figure 7a

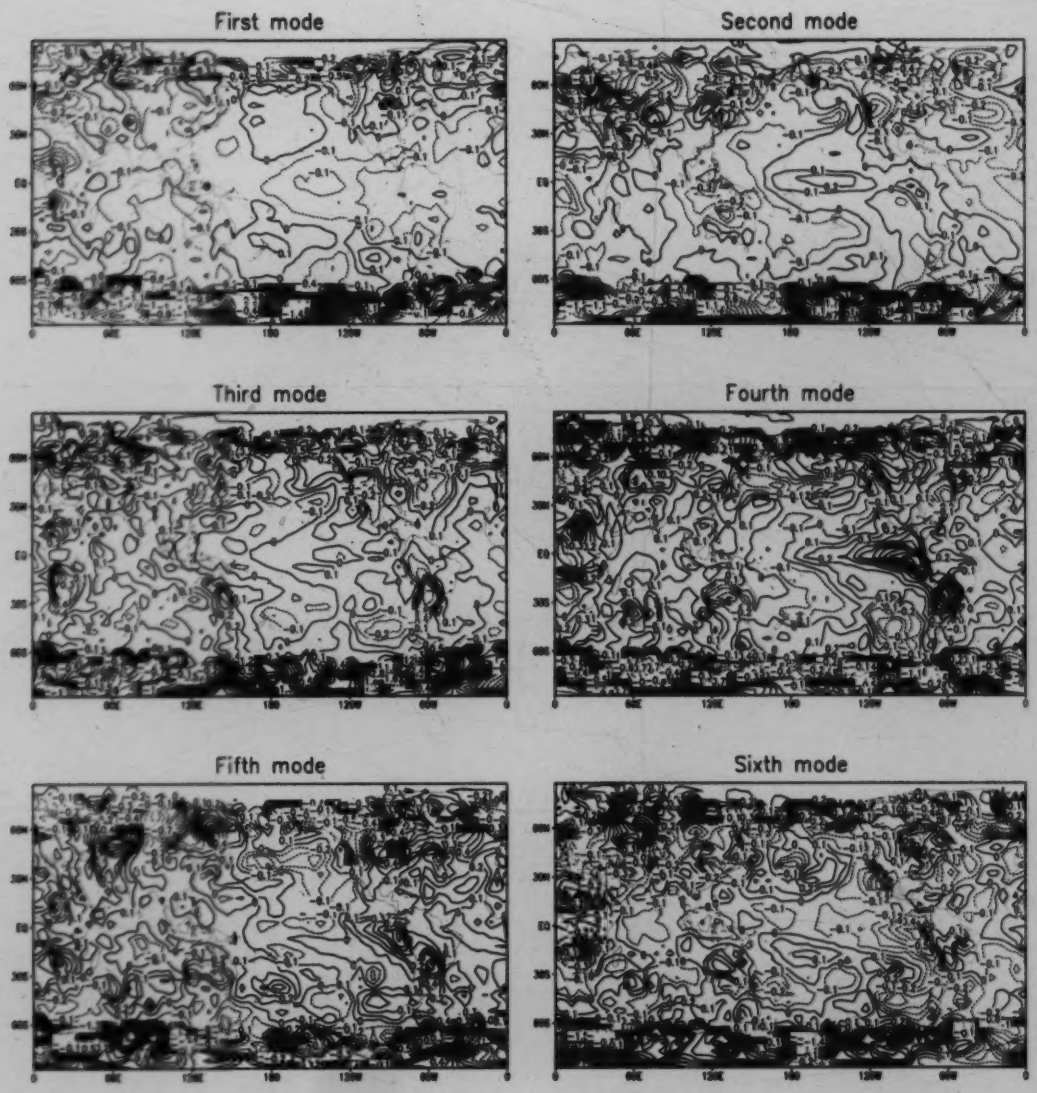

First mode
Second mode
Third mode
Fourth mode
Fifth mode
Sixth mode

Figure 7b

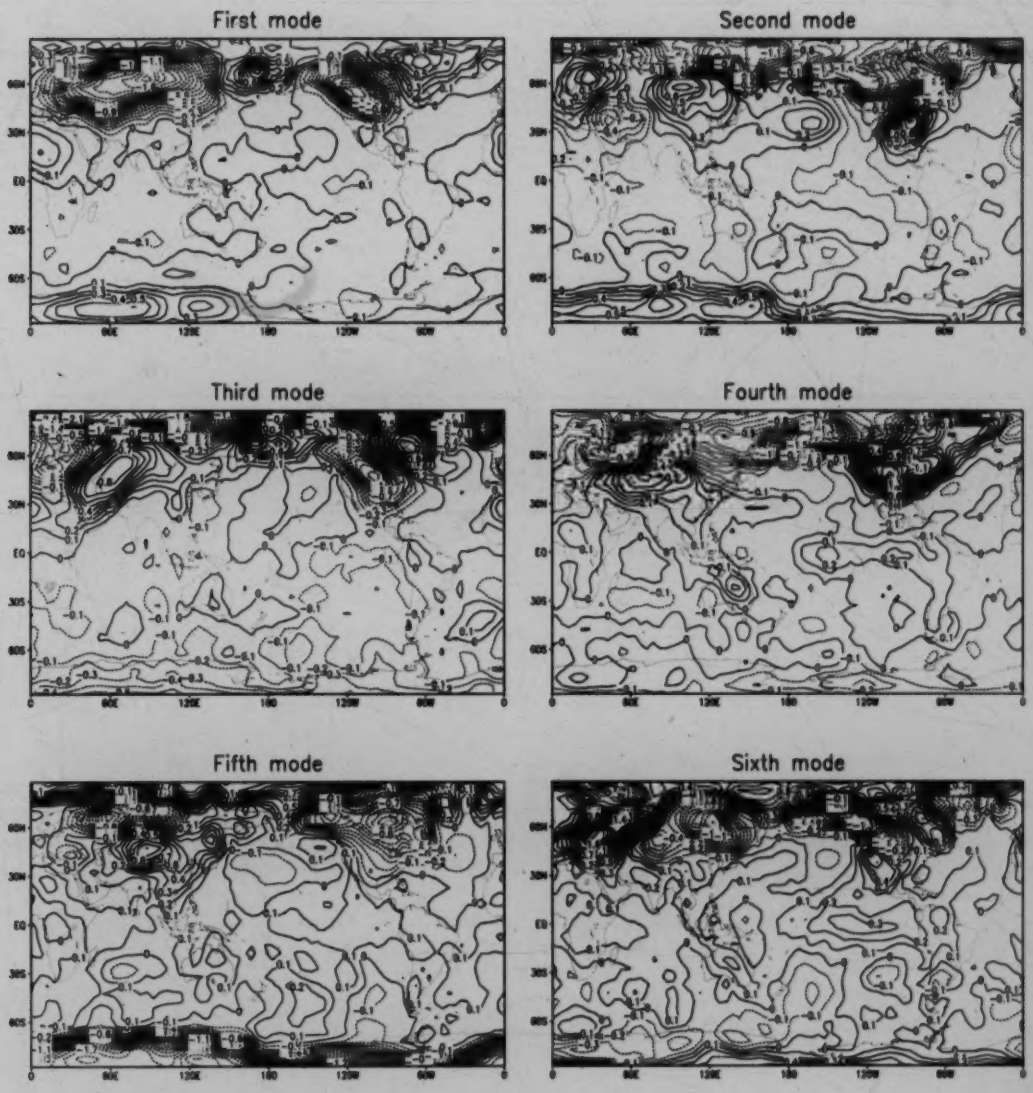

First mode
Second mode
Third mode
Fourth mode
Fifth mode
Sixth mode

Figure 8a

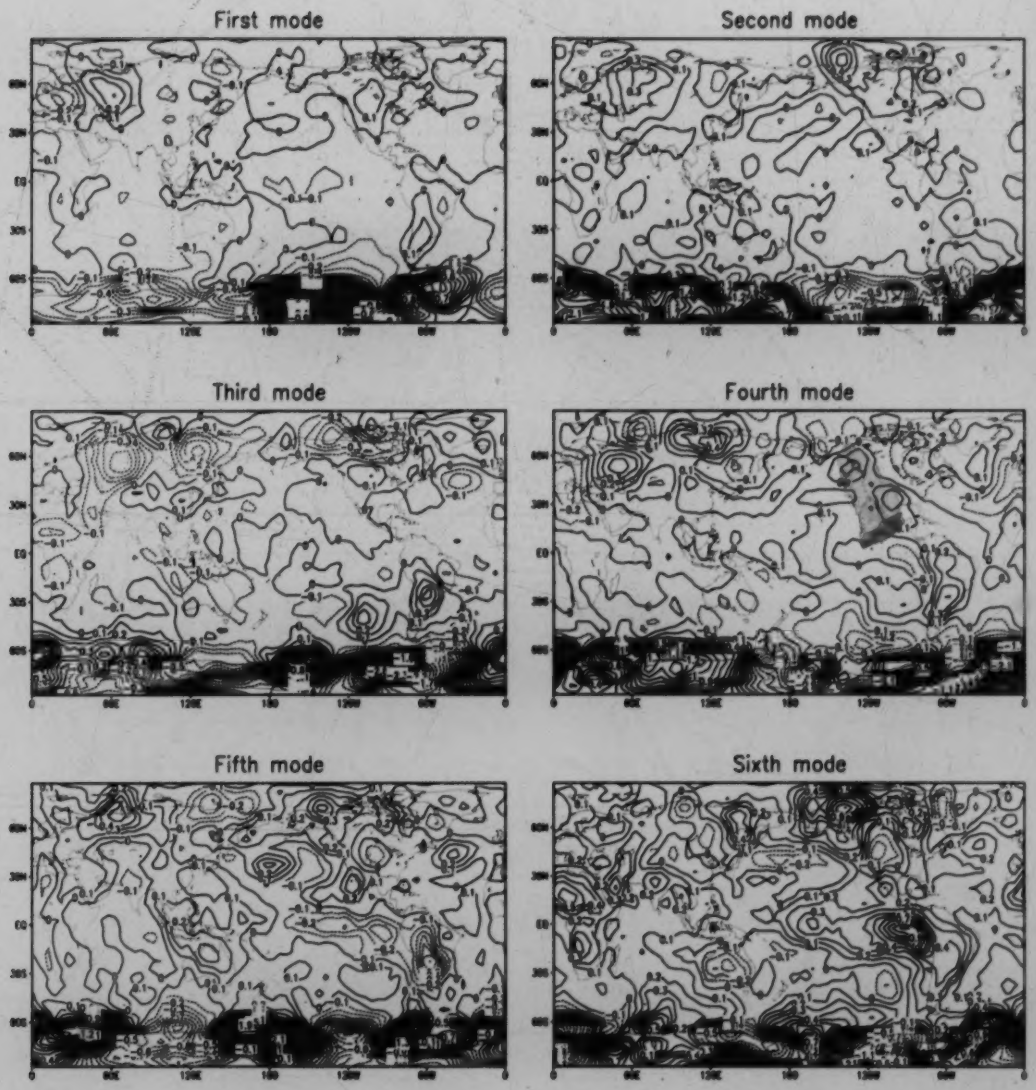
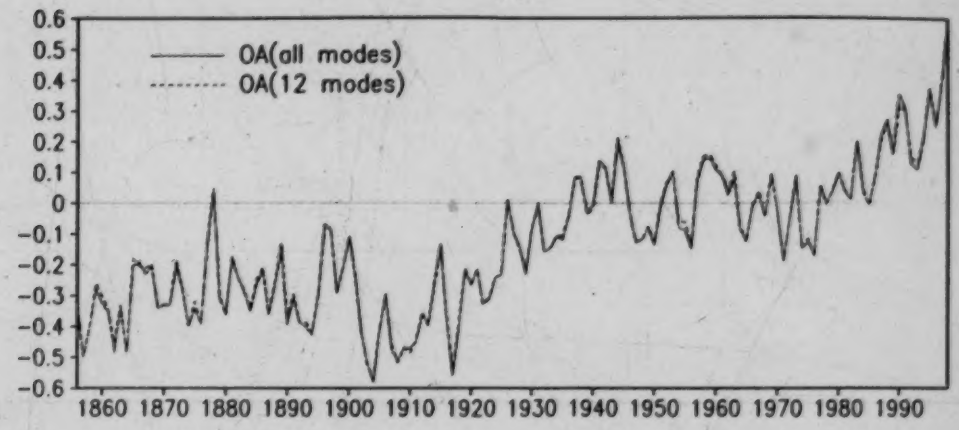
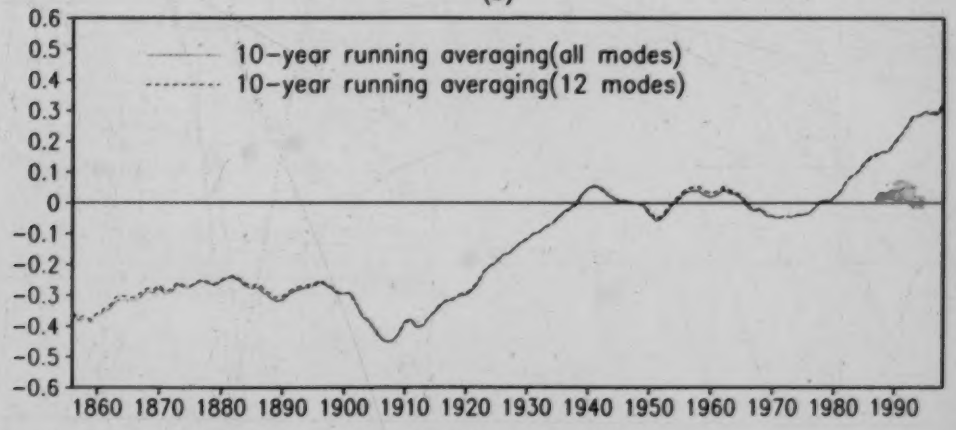

First mode
Second mode
Third mode
Fourth mode
Fifth mode
Sixth mode

Figure 8b

(a)



(b)



(c) The difference between the two optimal averagings

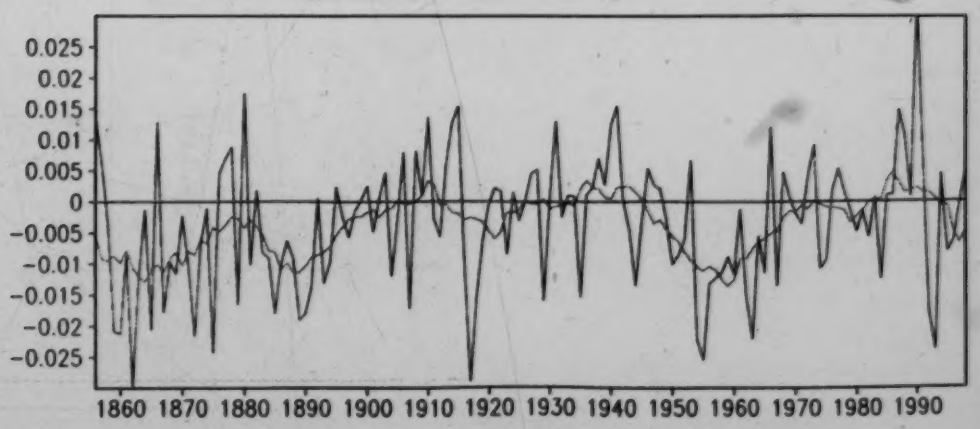


Figure 9

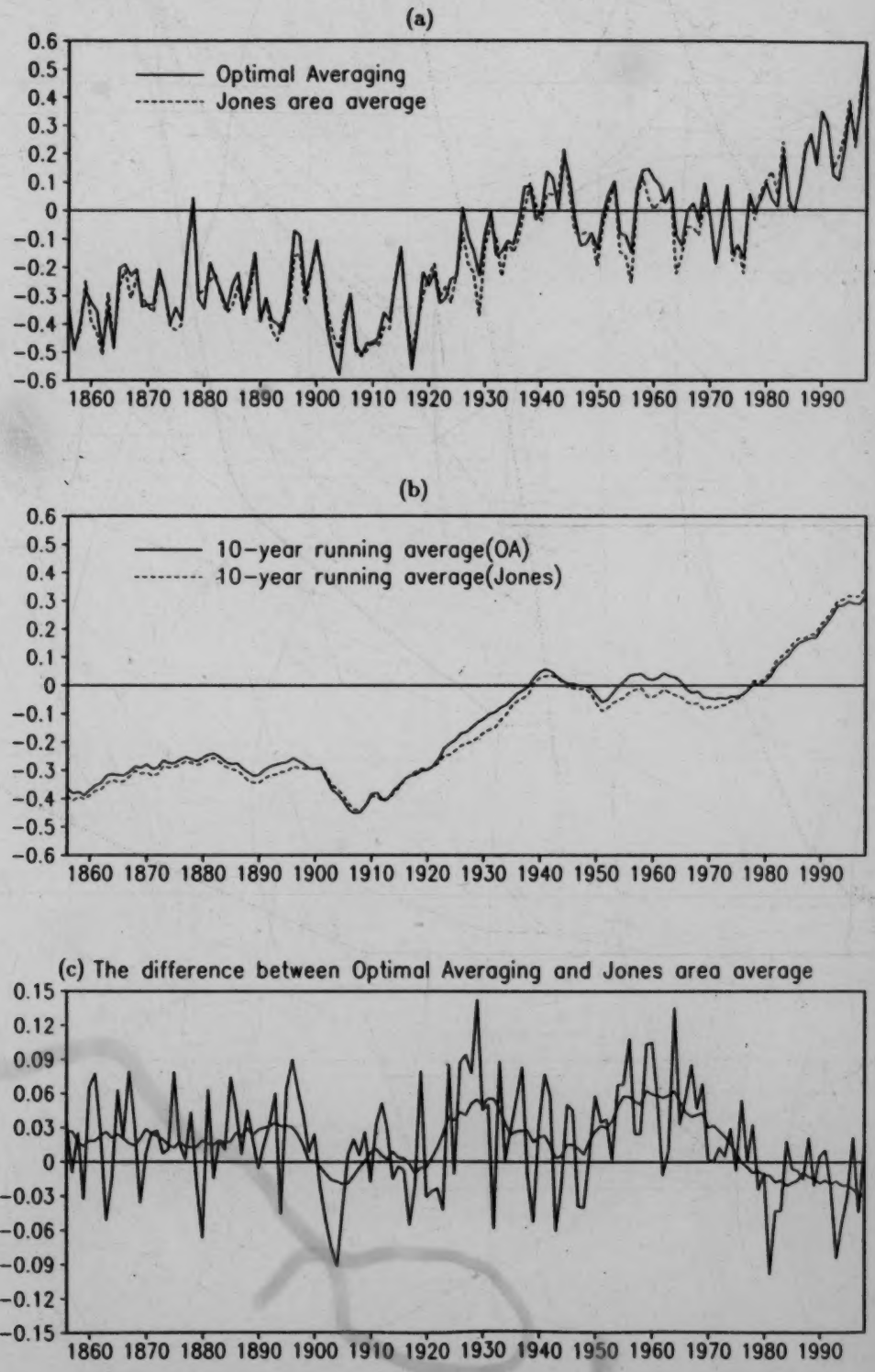

(a)
Optimal Averaging
Jones area average
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990
(b)
10−year running average(OA)
10−year running average(Jones)
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990
(c) The difference between Optimal Averaging and Jones area average
0.15
0.12
0.09
0.06
0.03
0
−0.03
−0.06
−0.09
−0.12
−0.15
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990

Figure 10

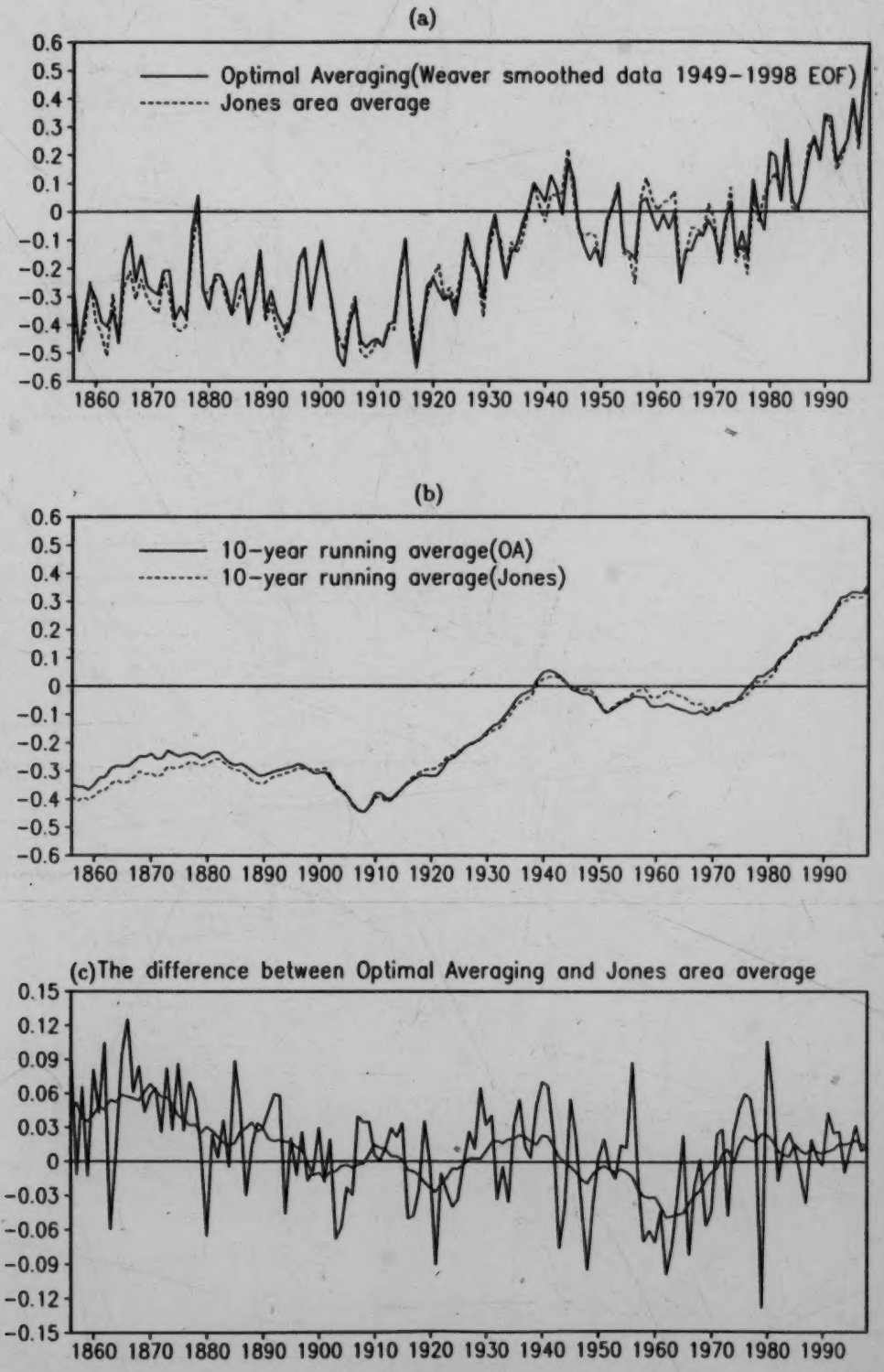

(a)
Optimal Averaging(Weaver smoothed data 1949-1998 EOF)
Jones area average
(b)
10-year running average(OA)
10-year running average(Jones)
(c)The difference between Optimal Averaging and Jones area average
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990

Figure 11

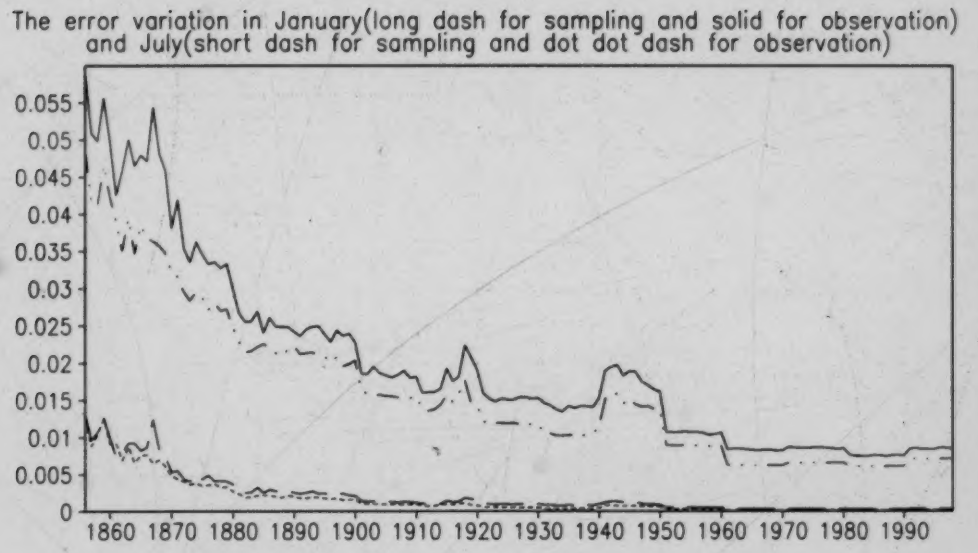

The error variation in January(long dash for sampling and solid for observation)
and July(short dash for sampling and dot dot dash for observation)
0.055
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990

Figure 12

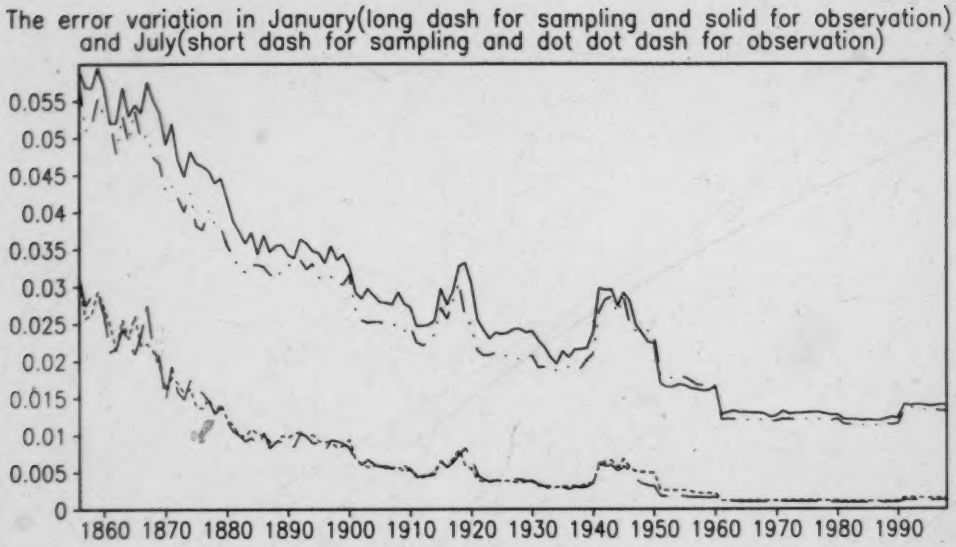

The error variation in January(long dash for sampling and solid for observation)
and July(short dash for sampling and dot dot dash for observation)
0.055
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990

Figure 13

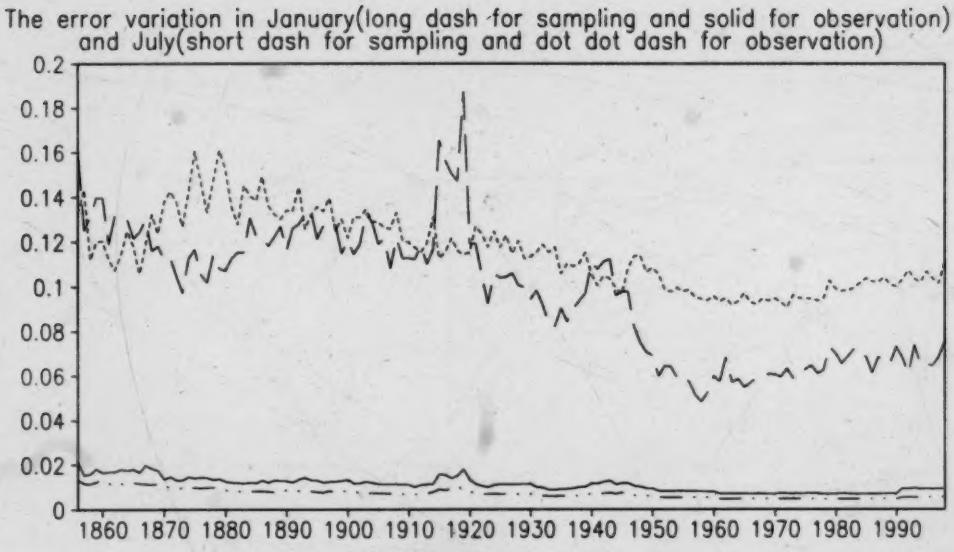

The error variation in January(long dash for sampling and solid for observation)
and July(short dash for sampling and dot dot dash for observation)
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990

Figure 14

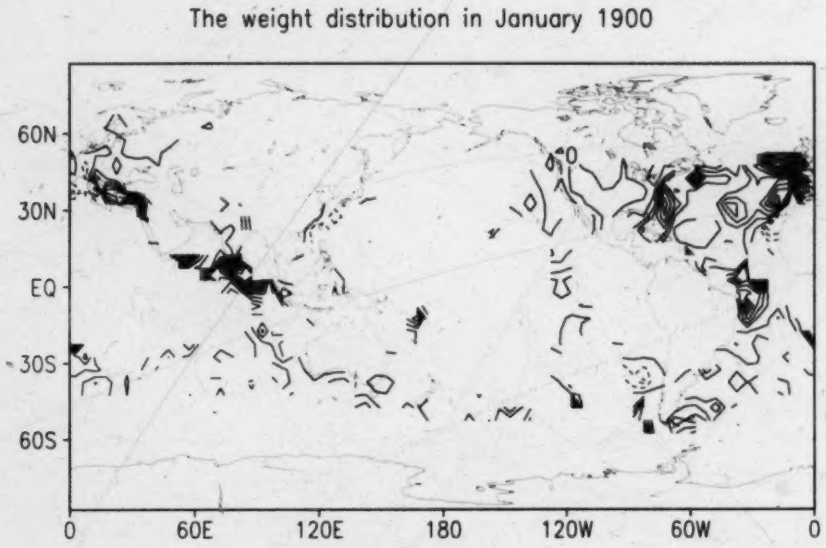

The weight distribution in January 1900
60N
30N
EQ
30S
60S
0
60E
120E
180
120W
60W
0

Figure 15a

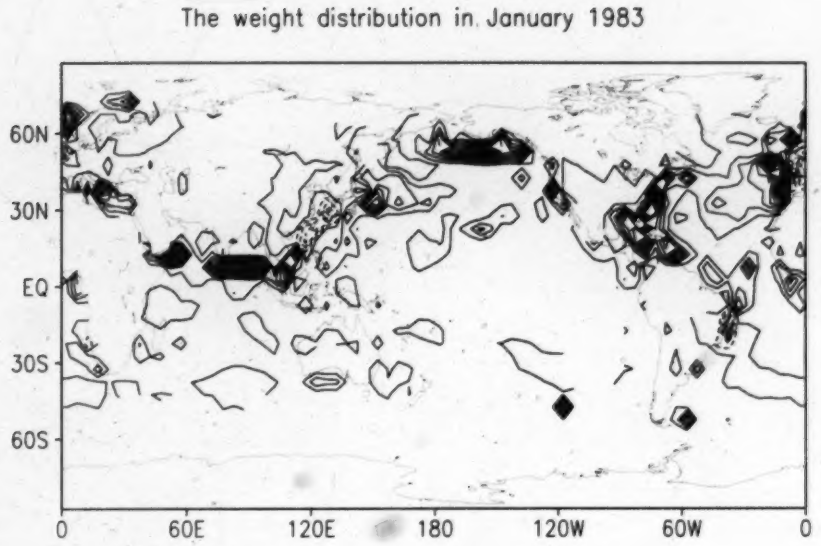

The weight distribution in January 1983
60N
30N
EQ
30S
60S
0
60E
120E
180
120W
60W
0

Figure 15b

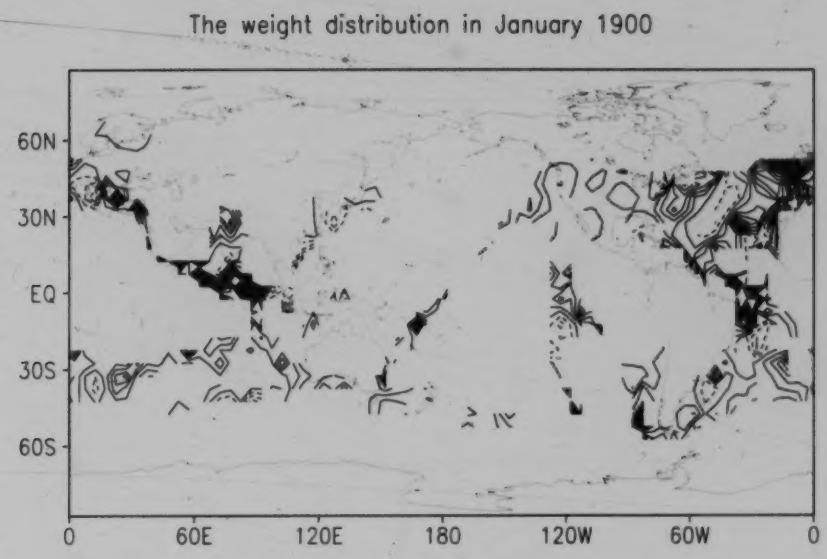

The weight distribution in January 1900
60N
30N
EQ
30S
60S
0
60E
120E
180
120W
60W
0

Figure 16a

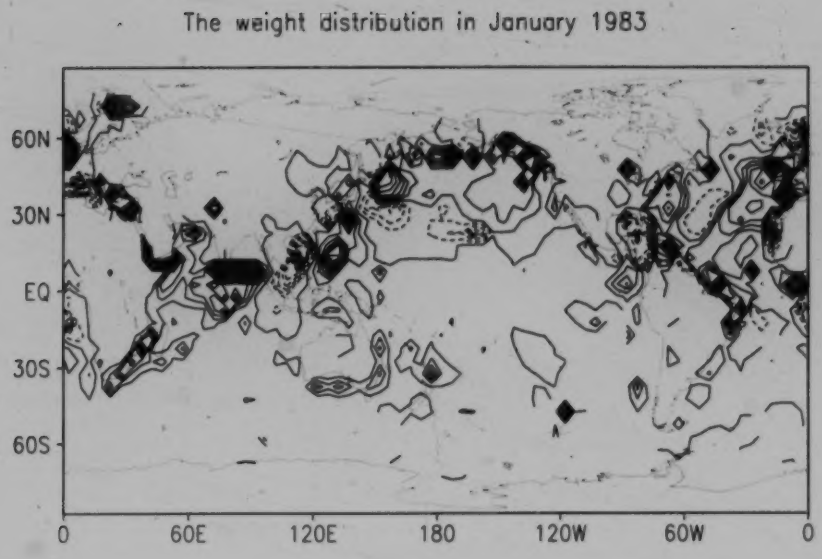

The weight distribution in January 1983
60N
30N
EQ
30S
60S
0
60E
120E
180
120W
60W
0

Figure 16b

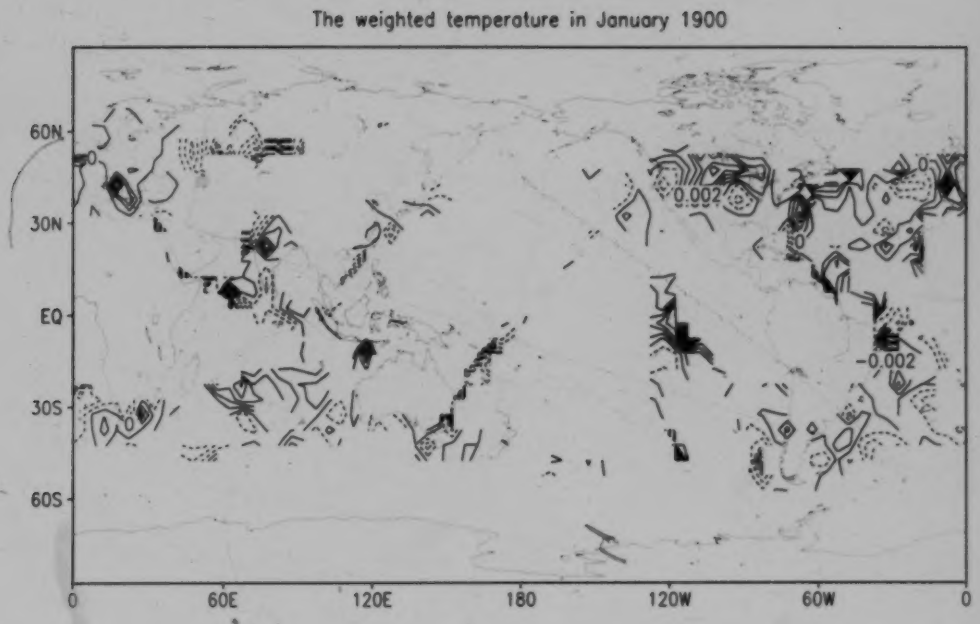

The weighted temperature in January 1900
60N
30N
EQ
30S
60S
0.002
0
-0.002
0
60E
120E
180
120W
60W
0

Figure 17a

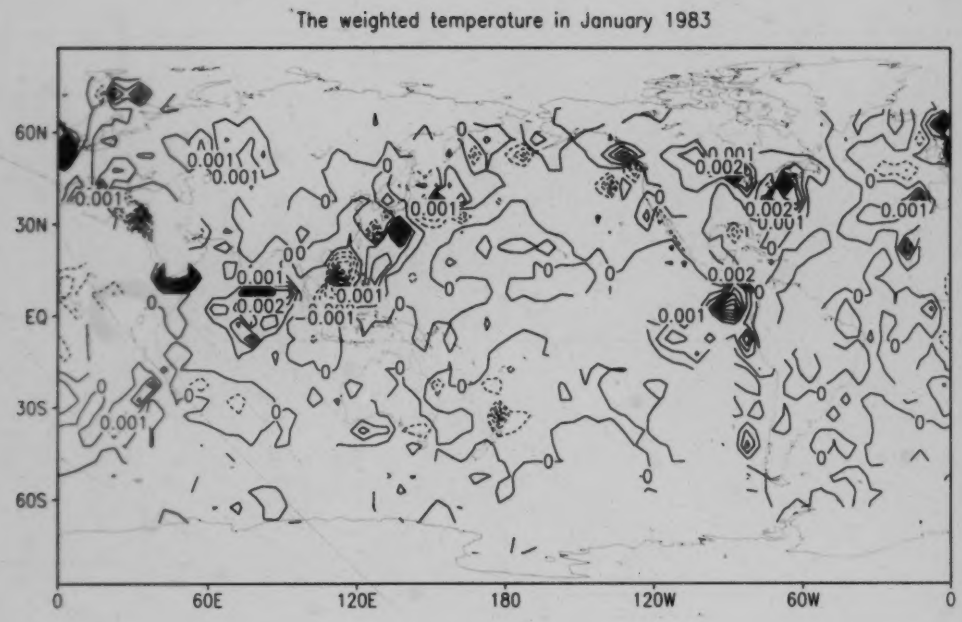

The weighted temperature in January 1983
60N
30N
EQ
30S
60S
0
60E
120E
180
120W
60W
0
0.001
0.002

Figure 17b

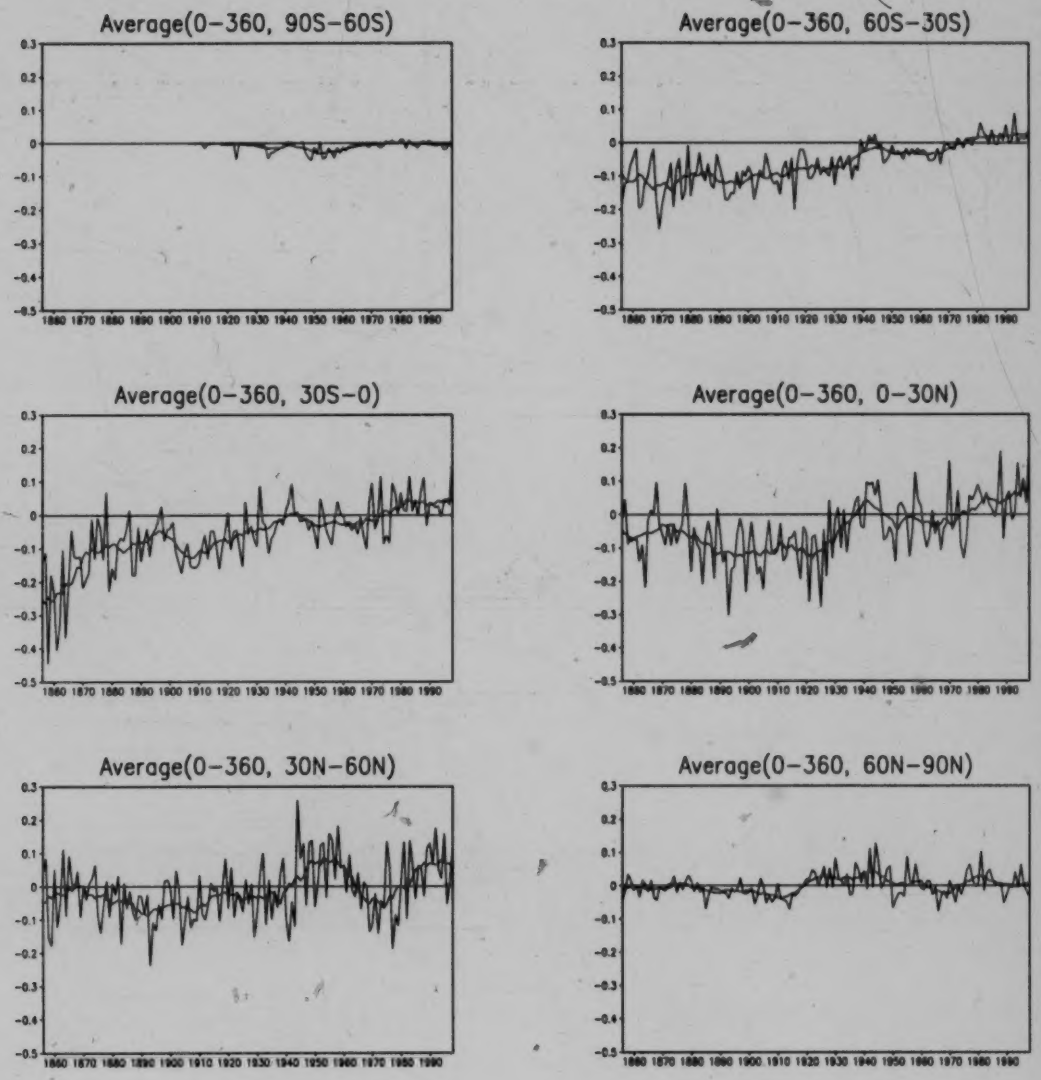

Average(0-360, 90S-60S)
Average(0-360, 60S-30S)
Average(0-360, 30S-0)
Average(0-360, 0-30N)
Average(0-360, 30N-60N)
Average(0-360, 60N-90N)

Figure 18

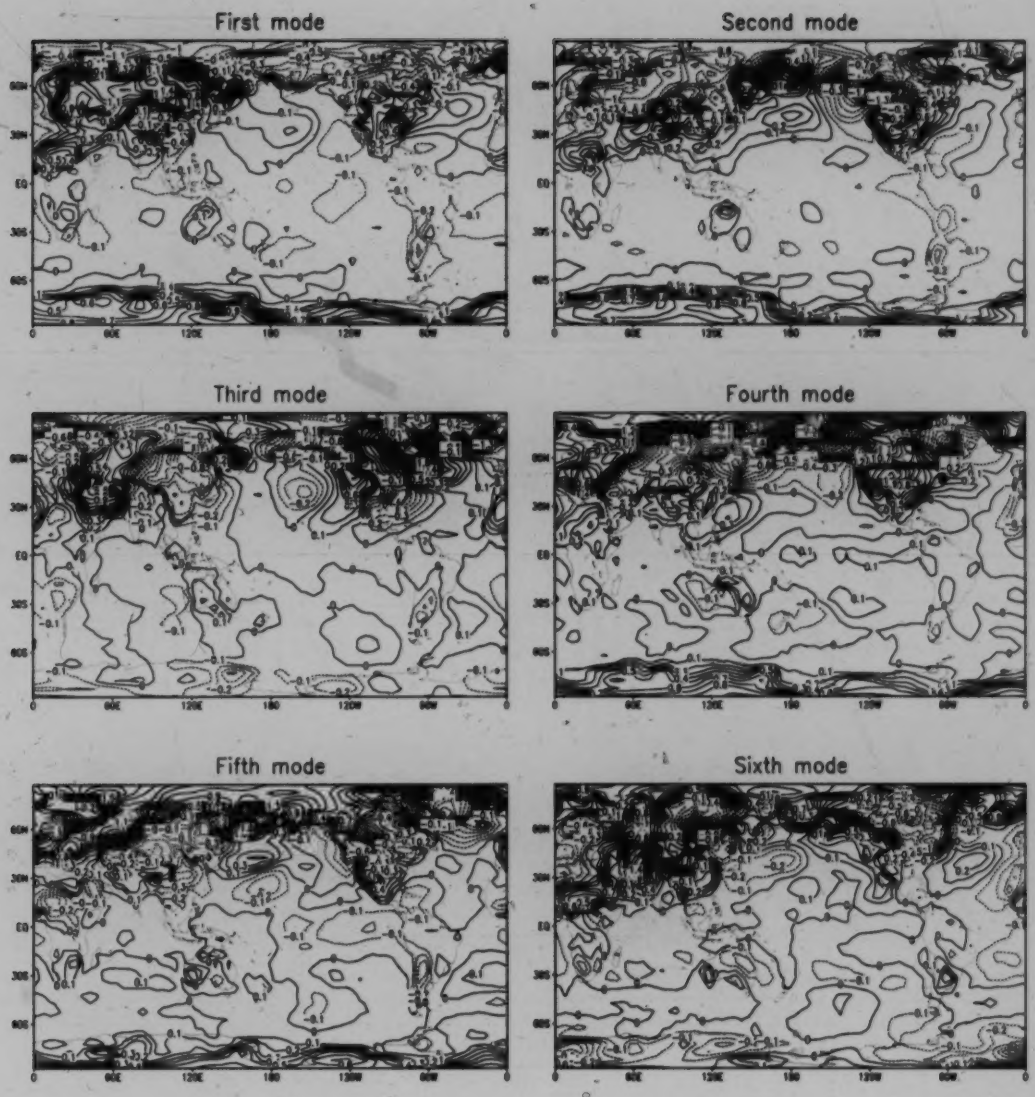

First mode
Second mode
Third mode
Fourth mode
Fifth mode
Sixth mode

Figure 19a

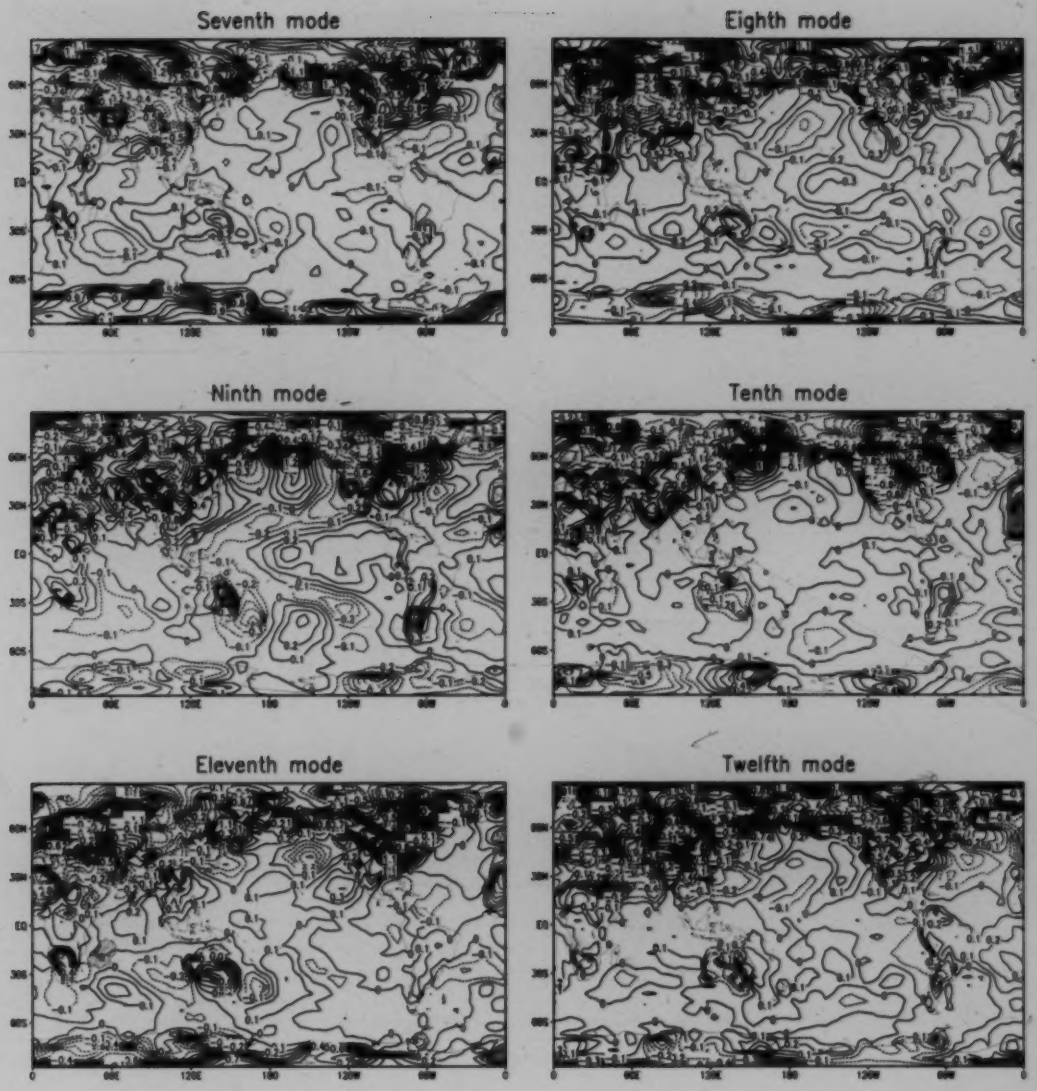

Seventh mode
Eighth mode
Ninth mode
Tenth mode
Eleventh mode
Twelfth mode

Figure 19b

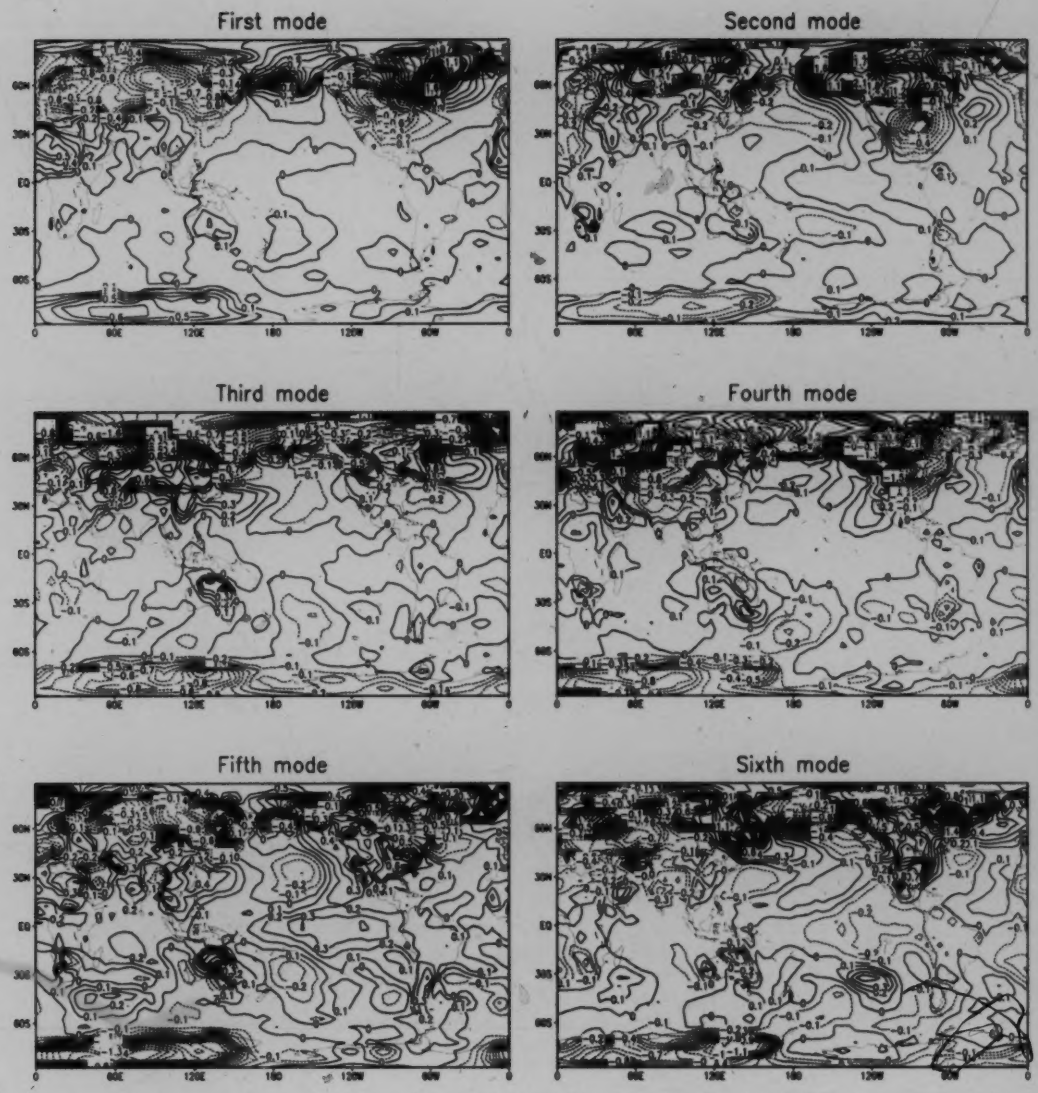

First mode
Second mode
Third mode
Fourth mode
Fifth mode
Sixth mode

Figure 20a

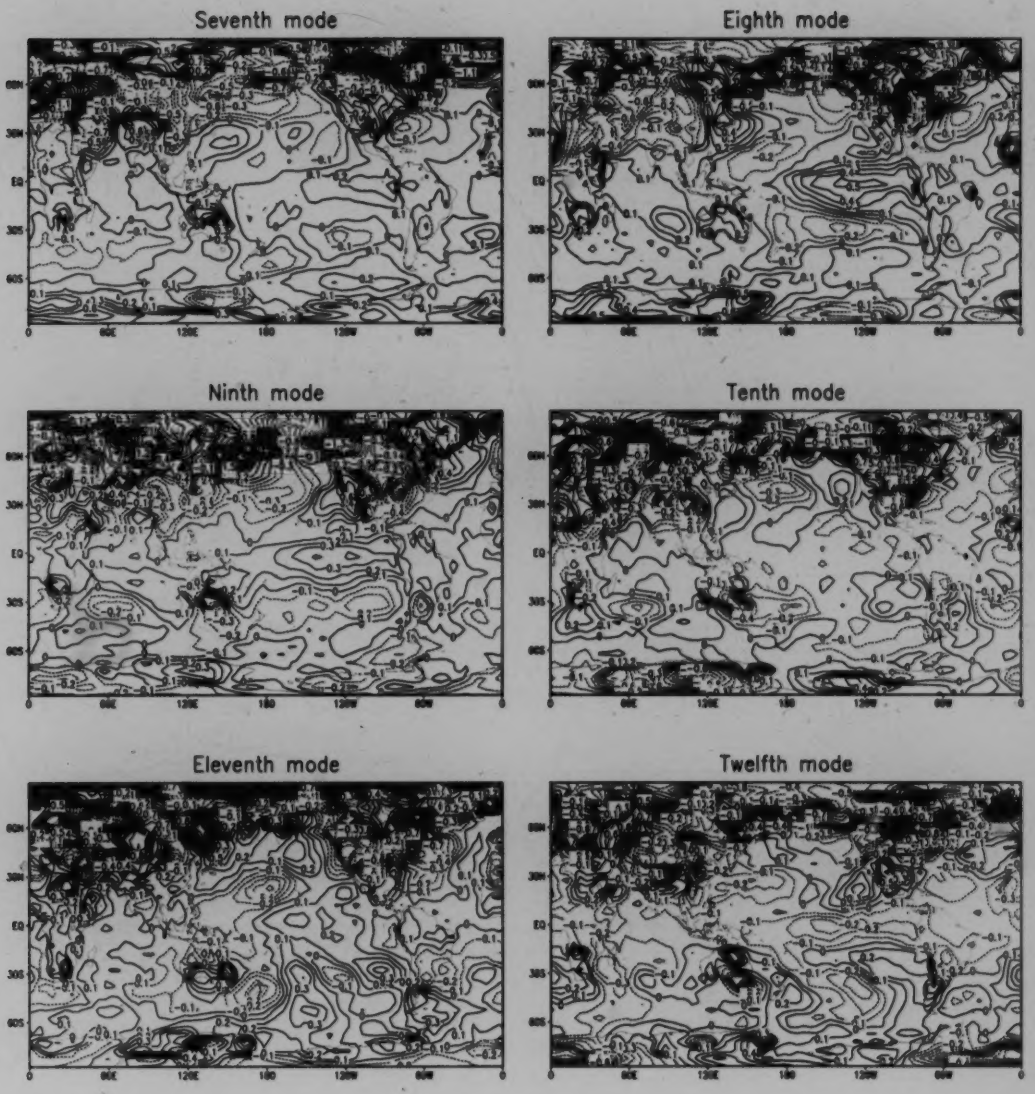

Seventh mode
Eighth mode
Ninth mode
Tenth mode
Eleventh mode
Twelfth mode

Figure 20b

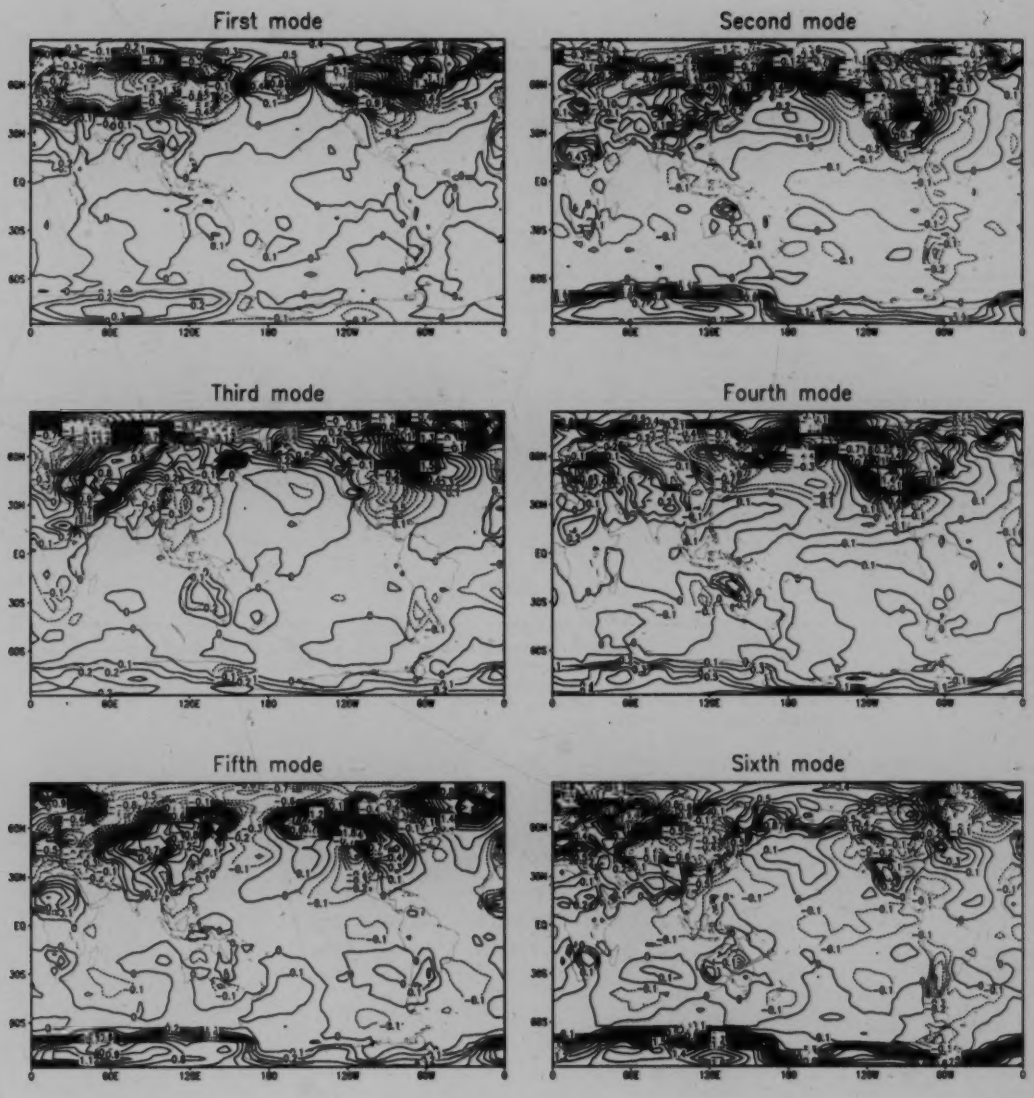

First mode
Second mode
Third mode
Fourth mode
Fifth mode
Sixth mode

Figure 21a

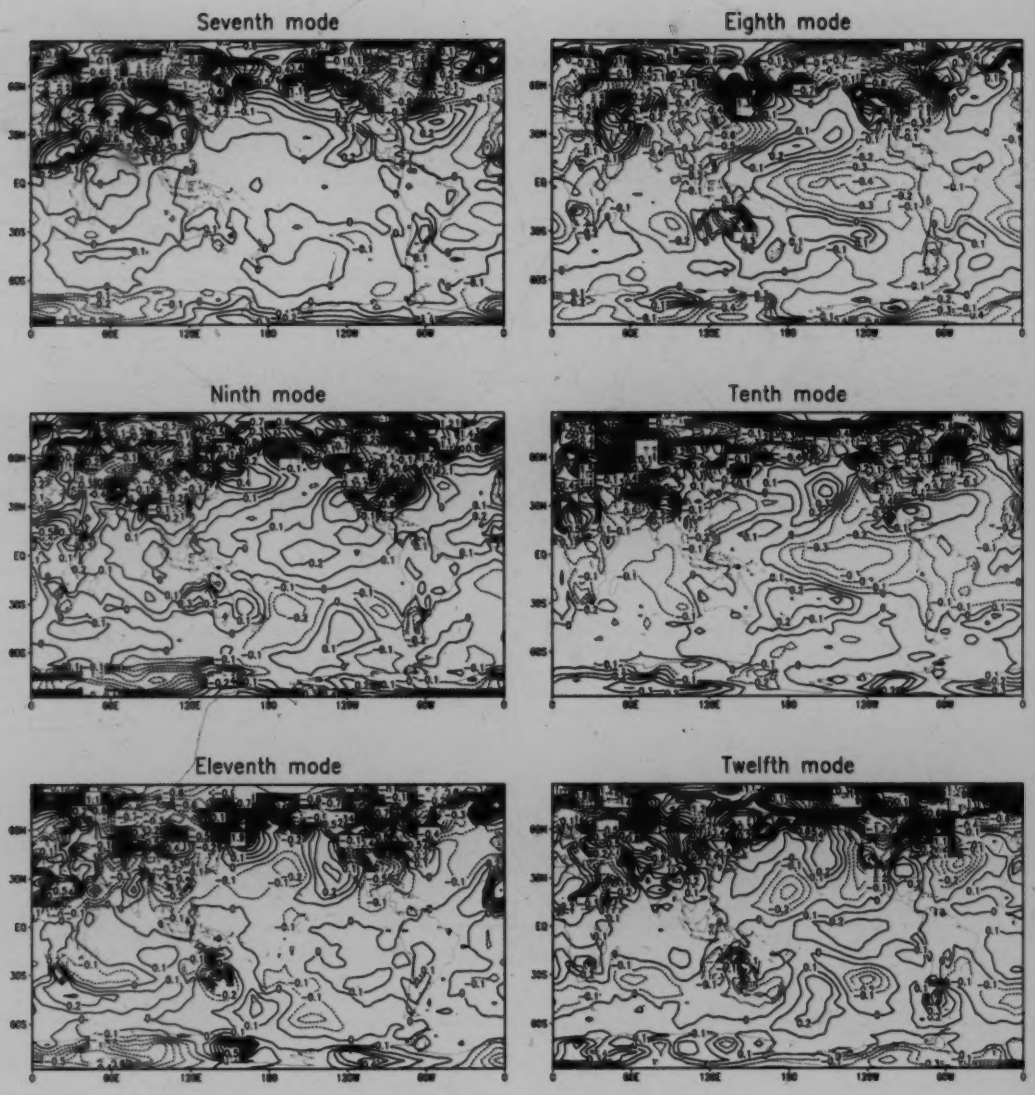

Seventh mode
Eighth mode
Ninth mode
Tenth mode
Eleventh mode
Twelfth mode

Figure 21b

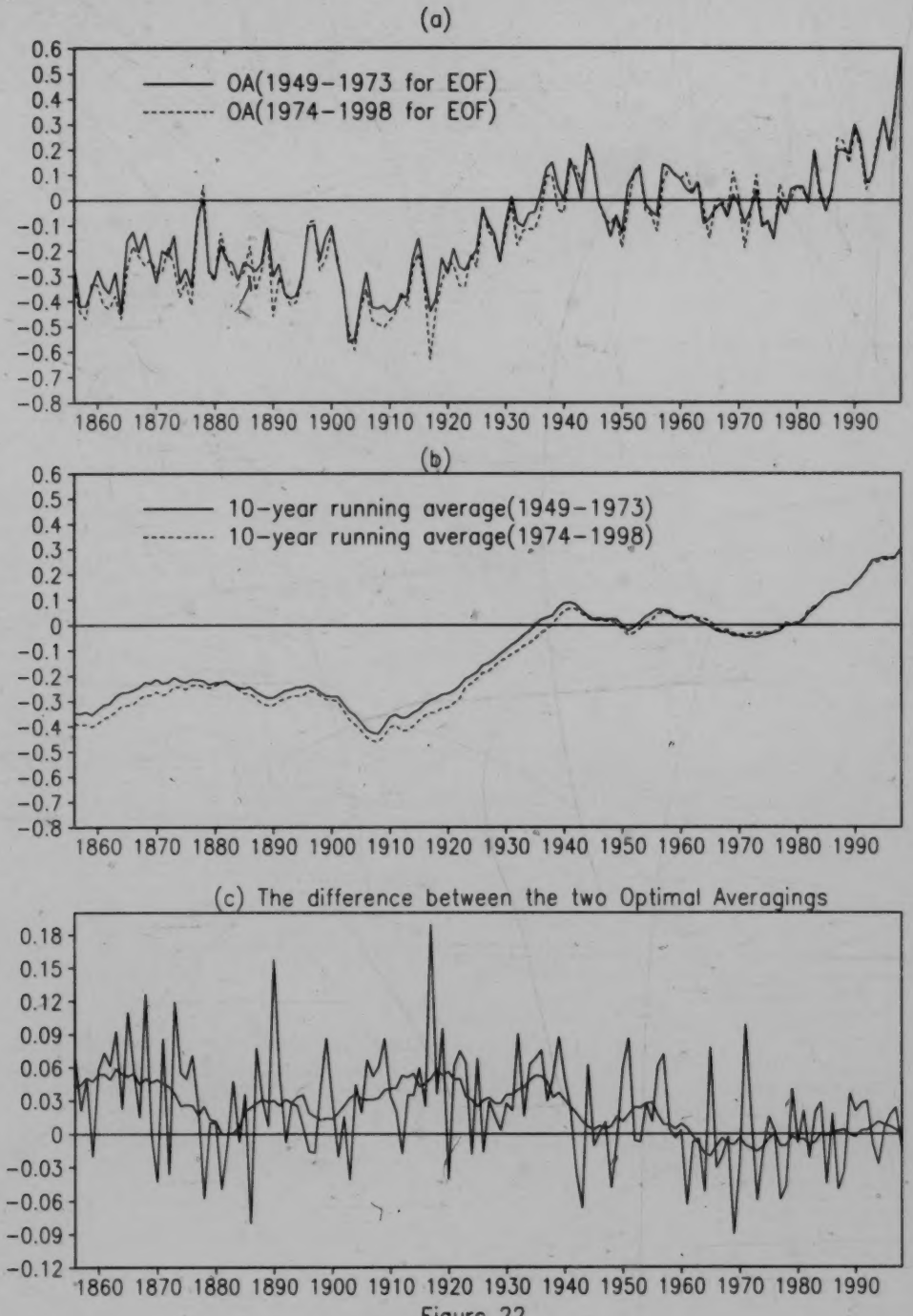

(a)
OA(1949–1973 for EOF)
OA(1974–1998 for EOF)
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
−0.7
−0.8
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990
(b)
10–year running average(1949–1973)
10–year running average(1974–1998)
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
−0.7
−0.8
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990
(c) The difference between the two Optimal Averagings
0.18
0.15
0.12
0.09
0.06
0.03
0
−0.03
−0.06
−0.09
−0.12
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990

Figure 22

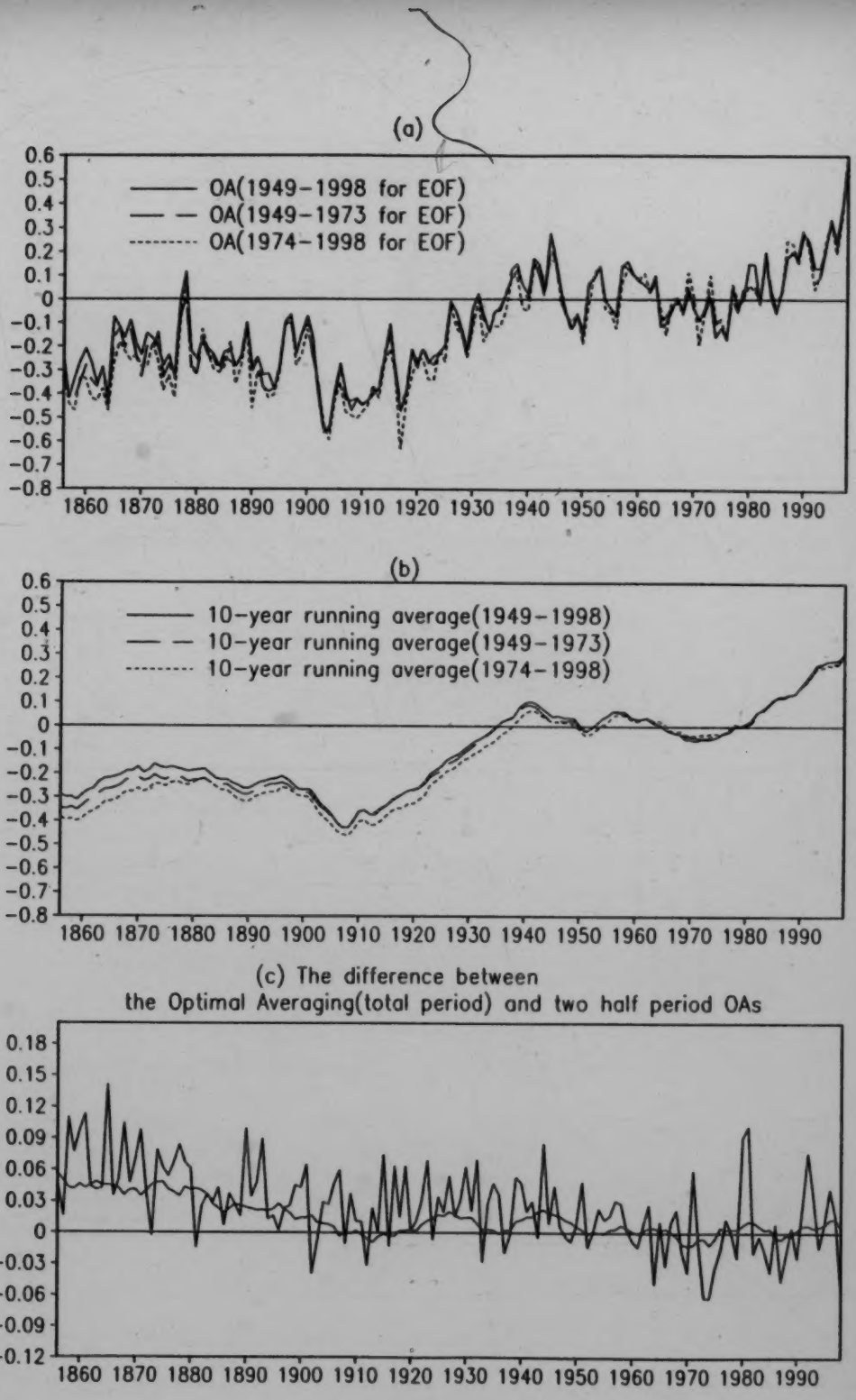

(a)
OA(1949-1998 for EOF)
OA(1949-1973 for EOF)
OA(1974-1998 for EOF)
0.6 0.5 0.4 0.3 0.2 0.1 0 -0.1 -0.2 -0.3 -0.4 -0.5 -0.6 -0.7 -0.8
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990
(b)
10-year running average(1949-1998)
10-year running average(1949-1973)
10-year running average(1974-1998)
0.6 0.5 0.4 0.3 0.2 0.1 0 -0.1 -0.2 -0.3 -0.4 -0.5 -0.6 -0.7 -0.8
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990
(c) The difference between
the Optimal Averaging(total period) and two half period OAs
0.18 0.15 0.12 0.09 0.06 0.03 0 -0.03 -0.06 -0.09 -0.12
1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990

Figure 23